

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

M

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>7</sup> :</b> <b>C07H 21/02, 19/04</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 00/46232</b> <b>(43) International Publication Date:</b> 10 August 2000 (10.08.00)
<b>(21) International Application Number:</b> PCT/US00/02897 <b>(22) International Filing Date:</b> 4 February 2000 (04.02.00)  <b>(30) Priority Data:</b> 60/118,721 5 February 1999 (05.02.99) US  <b>(71) Applicant:</b> INTEGRATED DNA TECHNOLOGIES, INC. [US/US]; 1710 Commercial Park, Coralville, IA 52241 (US).  <b>(72) Inventors:</b> BEHLKE, Mark, A.; 4 Fairview Knoll, Iowa City, IA 52240 (US). DEVOR, Eric, J.; 913 Wylde Green Road, Iowa City, IA 52246 (US). JAMES, Shale, A.; 1108 Wylde Green Road, Iowa City, IA 52246 (US). WALDER, Joseph, A.; 2709 West Sherwin, Chicago, IL 60645 (US).  <b>(74) Agents:</b> GORDON, Jennifer et al.; Pennie & Edmonds LLP, 1155 Avenue of the Americas, New York, NY 10036 (US).		<b>(81) Designated States:</b> AU, CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  <b>Published</b> <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
<b>(54) Title:</b> PRIMER EXTENSION METHODS FOR PRODUCTION OF HIGH SPECIFIC ACTIVITY NUCLEIC ACID PROBES  <b>(57) Abstract</b>  The present invention provides novel nucleic acid labeling techniques that generate nucleic acid probes with specific activities at least ten fold higher than the levels obtained using standard labeling methods. Specifically, the methods of the invention provide methods of producing nucleic probes that each comprises multiple labeled nucleotides. The methods can be used to generate RNA, DNA or hybrid probes. The invention also provides reaction mixtures and kits for the practice of the methods of the invention and compositions comprising the probes generated according to the methods of the invention.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LJ	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## PRIMER EXTENSION METHODS FOR PRODUCTION OF HIGH SPECIFIC ACTIVITY NUCLEIC ACID PROBES

This application claims priority, under 35 U.S.C. § 119(e), of United States  
5 Provisional Patent Application No. 60/118,721, filed February 5, 1999, which is  
incorporated by reference herein in its entirety.

### 1. INTRODUCTION

The present invention relates to nucleic acids, complexes, reaction mixtures, kits,  
10 and nucleic acid probe labeling techniques designed to increase the probe specific activity  
many-fold above the levels obtained using existing labeling techniques. Specifically, the  
present invention relates to methods of incorporating a plurality of labels into a single probe  
molecule. The present invention further relates to nucleic acid compositions labeled by  
such methods and to kits for performing the labeling methods of the invention.

15

### 2. BACKGROUND OF THE INVENTION

Oligonucleotides are DNA, RNA or DNA/RNA hybrid molecules that in general are  
relatively short, synthetic, and single stranded. Synthetic oligonucleotides are inexpensive,  
readily available, and can be made for any desired sequence. These molecules are often  
20 labeled with molecular tracers to follow their presence in various assays. Such labeled  
molecules are commonly referred to as "probes". The labeled oligonucleotide can be used  
as a hybridization probe to identify nucleic acids and studying the binding of proteins and  
other molecules to the probe. The results obtained using such assays depend in large part  
upon the specific activity of the probe. In general, probes of high specific activity that can  
25 be generated in a controllable and reproducible manner are most desirable.

Common nucleic acid labels currently include groups containing radioactive atoms  
such as  $^{32}\text{P}$  or  $^{35}\text{S}$ , biotin, fluorescent molecules (*e.g.* fluorescein and rhodamine derivatives)  
and enzymes whose activity is detectable, for example, as a result of catalyzing a reaction  
that produces a fluorescing or colorimetric reagent (*e.g.* horseradish peroxidase and alkaline  
30 phosphatase). In addition, nucleotide analogs conjugated to haptens (*e.g.* biotin,  
digoxigenin, *etc.*) may be used to label the nucleic acid. A hapten is subsequently  
recognized by a high affinity binding ligand (*e.g.*, streptavidin) or an antibody, that is  
conjugated to a fluorescent marker or enzyme reagent for detection of the nucleic acid.

Generally, polynucleotide kinase (PNK) is used to radiolabel oligonucleotides with  
35  $\gamma\text{-}^{32}\text{P}\text{-ATP}$ , in particular at the 5' phosphate. If there is no phosphate group present on the  
5' end of a nucleic acid molecule, PNK can add a radiolabeled phosphate to the free 5'

hydroxyl group of the nucleic acid. Alternatively, if a 5' phosphate is present, PNK can remove the non-radioactive 5' phosphate and replace it with a radiolabeled phosphate in an exchange reaction. However, in either an addition or exchange reaction, the resulting labeled nucleic acid has only a single  $^{32}\text{P}$  label per nucleic acid molecule. No more than one  $^{32}\text{P}$  label can be introduced per nucleic acid molecule by PNK.

The use of PNK is also limited, moreover, because the enzyme is not completely efficient in the addition reaction and even less efficient in the exchange reaction. This low efficiency can result in variable incorporation of radiolabel and a diminished specific activity measured in counts per minute (cpm) per micromole of DNA or RNA substrate.

Because of the inefficiency of the PNK, the enzyme reaction results leaves residual unlabeled nucleic acids. These unlabeled ("cold") nucleic acids cannot be purified away from the radiolabeled ("hot") nucleic acids because they have nearly identical chemical and physical properties. The labeled nucleic acid differs from the unlabeled molecule by the presence of only one additional phosphate group. The cold nucleic acid will then compete with the hot nucleic acid when the labeled nucleic acid is used in a practical application, e.g., binding to a target molecule, diminishing the probe's effectiveness.

In addition, because the unlabeled nucleic acid cannot be removed, the precise specific activity of the product generally cannot be determined. Usually, the quantity of oligonucleotide labeled is too small to measure the amount of product directly recovered from the reaction, e.g., by U.V. absorbance. Therefore, specific activity can only be estimated, either based on an assumption of the efficiency of recovery of the labeled nucleic acid or the recovery of unincorporated  $\alpha\text{-}^{32}\text{P}\text{-ATP}$ .

Nucleic acid molecules can also be labeled on their free 3' hydroxyl group with a radiolabeled nucleotide, e.g.,  $\alpha\text{-}^{32}\text{P}\text{-dATP}$ , using the enzyme Terminal Deoxynucleotidyl Transferase (TdT). As discussed below, TdT labeling, however, exhibits many of the same problems as PNK labeling. In particular, TdT does not efficiently label DNA, and, furthermore, the labeled nucleic acids cannot be separated from the unlabeled nucleic acids, reducing the efficiency of the method in which the labeled nucleic acid is being used.

Biologically, TdT is involved in increasing antigen receptor diversity by adding a variable number of nucleotides into chromosomal break points during the genetic rearrangement of the variable regions of immature T lymphocyte and B lymphocyte receptors (Gilfillan *et al.*, 1995, Imm. Reviews 148:201). TdT adds an average of four nucleotides to chromosomal break points *in vivo*; however, the number of nucleotides added is extremely variable and highly dependent on local conditions (Rock *et al.*, 1994, J. Exp. Med., 179:323). TdT also adds a variable number of nucleotides *in vitro* to the free 3' end of a nucleic acid probe, which makes it difficult to accurately determine the specific activity

of a probe or to produce the probe in a reproducible manner. Invariably, a mixture of species is formed from which it is very difficult to isolate a unique product. Thus, this method results in a heterogeneous reaction product ranging from unlabeled to hundred-fold labeled species. Further, TdT-labeled probes with a tail that is longer than 10 or 20 residues can produce high backgrounds when hybridized to complex nucleic acid samples as a result of binding to poly A tails of mRNAs or repeats in genomic DNA.

A number of schemes have been employed to increase the specific activity and sensitivity of oligonucleotide probes without compromising their favorable hybridization properties. Ullrich *et al.* (1984, *Nature* 309:418-425) and Studencki and Wallace (1984, *DNA* 3:7-15) described methods in which two overlapping sequence-specific oligonucleotides are labeled in a primer extension reaction to a specific activity that can exceed  $10 \times 10^9$  CPM/ $\mu$ g. While effective, these methods result in a double stranded product, require two sequence-specific custom oligonucleotides, and use two to four radiolabeled substrate deoxynucleotide triphosphates, making such approaches more expensive and less convenient than labeling with PNK. Oligonucleotides having a dendrimeric structure at the 5'-end can be synthesized and labeled to high specific activity using PNK (Shchepinov *et al.*, 1997, *Nucleic Acids Res.* 25:4447-4454). However these reagents are costly and are not readily available.

Both RNA and DNA polymerase-based techniques that incorporate labeled nucleotides into nucleic acids are well known in the art (see, *e.g.*, Ausubel *et al.*, 1992, "Short Protocols in Molecular Biology", John Wiley & Sons, New York). For example, nick translation methods allow the preparation of uniformly labeled DNA probes for a variety of assays. Small breakages are introduced into a double stranded DNA of interest (for example by treating with DNase I), and the gaps are "filled in" using DNA polymerase in the presence of radiolabeled or non-isotopically labeled nucleotides. Alternatively, labeled probes are generated by random priming. Briefly, a double-stranded DNA of interest is denatured and the individual strands replicated in the presence of the Klenow fragment of *E. coli* DNA polymerase I and labeled nucleotides, using random hexamers as primers for the extension step. Alternatively, a DNA which is to be used as a probe can be generated by cutting a parent DNA with a restriction enzymes that generates a 5' overhang, and the overhang filled in using the Klenow fragment of *E. coli* DNA polymerase I and labeled nucleotides. Yet another method of producing a labeled DNA probe is by generating the probe by polymerase chain reaction (PCR) in the presence of labeled nucleotides.

Generally, labeled RNA probes can be produced by carrying out the transcription reaction that generates the RNA in the presence of labeled nucleotides. Such transcription

reactions most commonly utilize bacteriophage RNA polymerases, such as SP6, T3, and T7.

Another scheme for labeling oligonucleotides proposed by Bahl *et al.* (U.S. Patent No. 5,538,872), involves a "bridging" technique that ligates a target molecule to a signal molecule via a "bridge" molecule. The signal molecule shows no complementarity towards the target molecule, and comprises a unique 4 or more nucleotide stretch at one end and is labeled at the other end. In addition to the sequence of interest, the target molecule comprises a unique 4 or more nucleotide stretch at the end opposite to that at which the unique sequence of the signal molecule is located. The signal and target molecules are ligated by a "bridge" molecule, consisting of two specific 4 or more nucleotide sequences, one complimentary to the unique nucleotide sequence in the target molecule, and the other complimentary to the unique nucleotide sequence in the signal molecule. The target molecule, the signal molecule and the bridging molecules are allowed to hybridize, allowing the target and signal molecules to align in a directional manner. The signal and target molecules are ligated together and the result is a double stranded oligonucleotide complex that can be left as is or disassociated into the bridge and a completed probe consisting of the target and signal molecules.

Another problem associated with conventional template-based labeling methods arises when the labeled probe must be separated from the template strand. Probe techniques have been attempted to resolve this problem. For example, Dattagupta *et al.*, U.S. Patent No. 4,808,520, proposed a probe that is labeled by hybridizing two oligonucleotides of equal length that have regions of mutual complementarity located at the 3' ends of the oligonucleotides. One overhang is filled in with labeled nucleotides to generate a labeled probe using the other strand as a template. The template strand is not extended because it has a different nucleotide composition from the strand that is labeled, and only nucleotides required to generate the probe are added to the labeling reaction. The probe strand can then be separated from the template strand based on the differing sizes of the strands. The drawback to this method is that the template strand is designed in a sequence specific manner that depends on the oligonucleotide that is being labeled.

Clearly, therefore, a need exists for methods whereby highly purified labeled nucleic acids of high specific activity can be produced in a controllable, reproducible manner.

### 3. SUMMARY OF THE INVENTION

The present invention provides nucleic acids, complexes, reaction mixtures, kits, and methods for labeling nucleic acids that generate high specific activity probes. The probes produced by the methods of the invention have specific activities that are many-fold higher than the specific activities obtained using existing labeling techniques. The present invention further relates to nucleic acid compositions produced by the methods of the invention and to kits for performing the labeling methods of the invention.

The nucleic acid labeling method of the present invention generates labeled nucleic acid probes, including but not limited to radiolabeled nucleic acid probes, with at least 10-fold higher specific activity than can be obtained by traditional 5' end-labeling with PNK. Unlike other methods of radioactive tailing (*i.e.*, TdT), the methods of the invention produce a single species of both known length and known specific activity. The reaction is highly efficient and approximately 90% of probe molecules are routinely labeled.

Specifically, the invention provides a method of labeling a nucleic acid molecule, comprising the steps of hybridizing a first nucleic acid, the Template Nucleic Acid, to a second nucleic acid, the Substrate Nucleic Acid. The Template Nucleic Acid comprises, from 3' to 5', a Substrate Hybridization Domain and a Signal Template Domain, wherein the Substrate Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides and the Signal Template Domain comprises a sequence of about 5 to about 100 nucleotides. The Substrate Nucleic Acid comprises, from 3' to 5', a Template Hybridization Domain and a Target binding domain, wherein the Template Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides, is not detectably labeled, and shows complementarity toward and is hybridizable to the Substrate Hybridization Domain of the Template Nucleic Acid, and the Target Binding Domain is not detectably labeled and comprises a nucleotide sequence heterologous to that of the Template Hybridization Domain of the Substrate Nucleic Acid. Upon hybridization, the Substrate Nucleic Acid is extended with a DNA polymerase in the presence of a labeled nucleotide to create a Signal Domain having a sequence which shows complementarity toward and is hybridizable to the Signal Template Domain of the Template Nucleic Acid. A Substrate Nucleic Acid comprising a Signal Domain is hereinafter called a Probe.

The labeled nucleic acid, *i.e.* the Probe, can also be purified from any remaining unlabeled nucleic acid, to obtain a species of very high, and precisely known, specific activity, for example by polyacrylamide gel electrophoresis ("PAGE").

The present invention particularly provides a method for obtaining labeled oligonucleotide probes comprising a plurality of labels per molecule. The plurality of labels can be of one or more species. The label can be a radiolabel, an enzyme, a hapten, or a

fluorescent molecule. The precise number of labels added incorporated into the Signal Domain can be adjusted within a wide range by changing the length or sequence of the Signal Template Domain of the Template Nucleic Acid strand or by altering the ratio of labeled nucleotides in the extension reaction.

- 5       The Template Hybridization Domain of the probe does not function as a target specific hybridization sequence, but rather serves as an attachment site for the Template Nucleic Acid. This sequence is preferably added to a nucleic acid at the time the nucleic acid is synthesized. In another embodiment, the Template Hybridization Domain is added by a fill-in reaction after hybridizing to a nucleic acid molecule containing a complementary  
10 overlap sequence.

- The Signal Template Domain serves as the template for the addition of labeled nucleotides. The preferred length of the Signal Template Domain is from about 5 to about 100 bases, with a length from about 10 to about 50 bases being preferred. Any sequence of bases can be employed. In a preferred embodiment, the sequence of the Signal Template  
15 Domain is oligo-dT. This provides, for example, a template for attachment of dATPs to the Substrate Nucleic Acid by primer extension catalyzed by a DNA polymerase, wherein the dATPs incorporate an appropriate label, *e.g.*,  $\alpha$ -<sup>32</sup>P-dATP.

- One skilled in the art will recognize that labeled nucleotides other than dATP can be used in combination with a template other than oligo-dT. The advantages of the  
20 preferred embodiment are that poly-dA in the Signal Domain has a lower melting temperature ( $T_m$ ) than sequences containing C or G residues, decreasing the likelihood of non-specific interactions, yet this sequence will not bind to the poly-A tail of mRNAs in hybridization reactions with cellular RNA.

- Prior to labeling, the Substrate Nucleic Acid and Template Nucleic Acid strands  
25 associate by hybridizing to each other through base pairing at their corresponding binding sites, *i.e.*, at the regions of complementarity, via the Template Hybridization Domain and the Substrate Hybridization Domain, respectively. The subsequent labeling reaction utilizes a DNA polymerase in a primer extension reaction to form a covalent attachment between the labeled nucleotides and the Substrate Nucleic Acid. DNA polymerases useful in the  
30 primer extension reaction include, but are not limited to, *E. coli* DNA polymerase I holoenzyme, a Klenow fragment of *E. coli* DNA polymerase I, T4 DNA polymerase, T7 DNA polymerase, T7 DNA polymerases, any DNA polymerase encoded by a thermophilic bacterium, including but not limited to *Taq* polymerase, and reverse transcriptase. The Klenow fragment of *E. coli* DNA Polymerase I is preferred. The enzymes can be native or  
35 recombinant and may be modified, *e.g.*, modified to remove certain activities such as, for



example, exonuclease activities. With respect to a Klenow fragment, for example, a Klenow fragment lacking the 3'-exonuclease activity of the enzyme can be utilized.

The present invention also provides a Probe nucleic acid comprising, from 3' to 5', a Signal Domain, a Template Hybridization Domain and a Target Binding Domain, wherein  
5 the Signal Domain comprises a sequence of about 5 to about 100 nucleotides of which at least two nucleotides are detectably labeled; the Template Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides, and is not detectably labeled; and the Target Binding Domain is not detectably labeled and comprises a nucleotide sequence heterologous to that of the Template Hybridization Domain, and is not detectably labeled.

10 In one embodiment, the nucleotides which comprise the Probe nucleic acid are deoxyribonucleotides. In another embodiment, the nucleotides which comprise the Signal Domain of the Probe are deoxyribonucleotides and the nucleotides which comprise the Template Hybridization Domain and the Target Binding Domain of the Probe are ribonucleotides. In certain embodiments of the invention, the Signal Domain of the Probe  
15 comprises a sequence of about 10 to about 50 nucleotides.

The present invention provides a Probe in which the Signal Domain is at least 50%, at least 70%, at least 90% or 100% homopolymeric.

The present invention provides a Probe in which at least one nucleotide comprises a label selected from the group consisting of:  $^{32}\text{P}$ ,  $^{33}\text{P}$ ,  $^{35}\text{S}$ , fluorescein, digoxigenin, biotin,  
20 Cy5, Cy3, and rhodamine.

The present invention provides a Probe in which the Template Hybridization Domain comprises a sequence of about 5 to about 10 nucleotides.

The present invention further provides a Probe in which at least 60%, at least 80% or 100% of the nucleotides of the Template Hybridization Domain comprise guanosine or  
25 cytidine or a combination thereof. In instances wherein 100% of the nucleotides of the Template Hybridization Domain comprise guanosine or cytidine, it is preferred that there be a combination of guanosine and cytidine present.

In certain embodiments of the present invention, at least 60% of the nucleotides of the Template Hybridization Domain of the Probe comprise guanosine or cytidine or a  
30 combination thereof, and the Signal Domain of the Probe is at least 50% homopolymeric.

The present invention also provides a Complex, comprising a Template Nucleic Acid comprising, from 3' to 5' a Substrate Hybridization Domain and a Signal Template Domain, wherein the Substrate Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides and the Signal Template Domain comprises a sequence of about 5 to  
35 about 100 nucleotides; and a Probe nucleic acid comprising from 3' to 5' a Signal Domain, a Template Hybridization Domain and a Target Binding Domain, wherein the Signal Domain

comprises a sequence of about 5 to about 100 nucleotides, which sequence shows complementarity toward and is hybridizable to the Signal Template Domain of the Template Nucleic Acid, and of which at least two nucleotides are detectably labeled; the Template Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides, is not detectably labeled, and shows complementarity toward and is hybridizable to the Substrate Hybridization Domain of the Template Nucleic Acid; the Target Binding Domain is not detectably labeled and comprises a nucleotide sequence heterologous to that of the Template Hybridization Domain of the Substrate Nucleic Acid.

The present invention provides a Complex in which the nucleotides which comprise the Probe or Template Nucleic Acid are deoxyribonucleotides. The present invention provides a Complex in which the nucleotides which comprise the Template Nucleic Acid are ribonucleotides. The present invention provides a Complex in which the nucleotides which comprise the Signal Domain of the Probe are deoxyribonucleotides and the nucleotides which comprise the Template Hybridization Domain and the Target Binding Domain of the Probe are ribonucleotides.

The present invention provides a Complex in which the Substrate Hybridization Domain is at the 3' end of the Template Nucleic Acid.

The present invention provides a Complex in which the Substrate Hybridization Domain of the Template Nucleic Acid comprises a sequence of about 5 to about 10 nucleotides.

The present invention provides a Complex in which the Substrate Hybridization Domain of the Template Nucleic Acid cannot be extended by a 5'-3' DNA polymerase.

The present invention provides a Complex in which the Substrate Hybridization Domain of the Template Nucleic Acid further comprises an extension of nucleotides at the 3' end of said Substrate Hybridization Domain, the extension having no complementarity to the Template Hybridization Domain of the Substrate Nucleic Acid.

the Substrate Hybridization Domain comprises a 3'-terminal modified nucleotide. The modification is selected from the group consisting of: a 3'-amino-modifier, a 2', 3'-dideoxynucleotide, a 3'-phosphate, and a modified 3'-phosphate group.

The present invention provides a Complex in which the Substrate Hybridization Domain comprises at least one nucleotide which comprises a modified cytidine, which nucleotide is selected from the group consisting of: C5-methyl-dC and C5-propynyl-dC.

The present invention further provides a Complex in which the Signal Domain of the Probe comprises a sequence of about 10 to about 50 nucleotides.

The present invention further provides a Complex in which the Signal Domain is at least 50%, at least 70%, at least 90% or 100% homopolymeric.

The present invention further provides a Complex in which the at least 60%, at least 80% or 100% of the nucleotides of the Template Hybridization Domain comprise guanosine or cytidine or a combination thereof.

5 The present invention yet further provides a Complex in which the at least 60% of the nucleotides of the Template Hybridization Domain comprise guanosine or cytidine or a combination thereof, and the Signal Domain is at least 50% homopolymeric.

10 The present invention further provides a reaction mixture for use in a process for the labeling of a nucleic acid molecule comprising a Template Nucleic Acid comprising, from 3' to 5', a Substrate Hybridization Domain and a Signal Template Domain, wherein the Substrate Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides, and the Signal Template Domain comprises a sequence of about 5 to about 100 nucleotides; and a Substrate Nucleic Acid comprising from 3' to 5', a Template Hybridization Domain and a Target Binding Domain, wherein the Template Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides, is not detectably labeled, and shows  
15 complementarity toward and is hybridizable to the Substrate Hybridization Domain of the Template Nucleic Acid; the Target Binding Domain is not detectably labeled and comprises a nucleotide sequence heterologous to that of the Template Hybridization Domain of the Substrate Nucleic Acid.

20 The present invention further provides a kit for labeling a nucleic acid molecule, comprising a reaction mixture comprising Template Nucleic Acid comprising, from 3' to 5', a Substrate Hybridization Domain and a Signal Template Domain, wherein the Substrate Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides, and the Signal Template Domain comprises a sequence of about 5 to about 100 nucleotides; a Substrate Nucleic Acid comprising from 3' to 5', a Template Hybridization Domain and a  
25 Target Binding Domain, wherein the Template Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides, is not detectably labeled, and shows complementarity toward and is hybridizable to the Substrate Hybridization Domain of the Template Nucleic Acid; the Target Binding Domain is not detectably labeled and comprises a nucleotide sequence heterologous to that of the Template Hybridization Domain of the Substrate  
30 Nucleic Acid; and a DNA polymerase.

The Probes of the invention can be used as either single stranded probes or, when complexed to Template Nucleic Acids, as partially double stranded nucleic acid probes. The double stranded region of such a probe will not interfere with specific binding of the unique sequence at the 5' end of the molecule. The Probes can be used in a variety of  
35 assays, including but not limited to electrophoretic mobility shift assays, protein binding

assays, Southern blots, Northern blots, Southwestern blots, dot blots, slot blots, colony and plaque lifts, and *in situ* hybridizations.

### 3.1 DEFINITIONS

5        Oligonucleotide: An oligonucleotide is single stranded nucleic acid molecule that is relatively short, generally no longer than about 150 nucleotides. An oligonucleotide can be a deoxyribonucleic acid, a ribonucleic acid, or a hybrid thereof.

Template Nucleic Acid: A Template Nucleic Acid is preferably an oligonucleotide, and comprises, from 3' to 5', a domain referred to herein as a "Substrate Hybridization  
10 Domain" or "Substrate Binding Domain" and a domain referred to herein as a "Signal Template Domain." The Substrate Hybridization Domain hybridizes to the Template Hybridization Domain of the Substrate Nucleic Acid, and is preferably at the 3'-terminus of the molecule. Upon hybridization of the Substrate Hybridization domain is the "Signal Template Domain", which upon hybridization of the Substrate Hybridization Domain to the  
15 Substrate Nucleic Acid, the Signal Template Domain forms a single stranded 5' overhang in a heteroduplex that can be used to extend the Substrate Nucleic Acid strand using a DNA polymerase and labeled nucleotides. The Template Nucleic Acid has a sequence independent from and heterologous to the sequence of the Target Binding Domain of the Substrate Nucleic Acid (see below), and can, therefore, be used for labeling any Substrate  
20 Nucleic Acid of interest. As such, this sequence is a "universal" sequence, and, therefore, can also be referred to as a "Universal Template Nucleic Acid."

Substrate Nucleic Acid: A nucleic acid, preferably an oligonucleotide, comprising from 3' to 5', a domain herein referred to as a "Template Hybridization Domain" or "Template Binding Domain" that is complementary to the Substrate Hybridization Domain  
25 of the Template Nucleic Acid and is heterologous to, *i.e.*, does not follow the Target Binding Domain in its native context, for example in the gene to which the Target Binding Domain corresponds; and a domain referred to herein as a "Target Binding Domain", which is a specific sequence for binding to a Target molecule of choice, *e.g.* for detection of a Target Nucleic Acid sequence or to test binding of a Target Protein to a particular DNA or  
30 RNA sequence in an electrophoretic mobility shift assay. After hybridization of the Substrate Nucleic Acid to the Template Nucleic Acid, the Template Hybridization Domain can serve as a primer for the extension of the Substrate Nucleic Acid via a DNA polymerase in the presence of at least one labeled nucleotide, with the Signal Template Domain of the Template Nucleic Acid functioning as a template for the extension reaction. The 3' domain  
35 of the Substrate Nucleic Acid and the molecule as a whole resulting from the extension are the "Signal Domain" and the "Probe," respectively.

**Complex:** As used herein, a "Complex" comprises a fully or partially hybridized pair of nucleic acids, for example a Template Nucleic Acid and a Substrate Nucleic Acid, wherein the Substrate Nucleic Acid is labeled. The Substrate Nucleic Acid in such a complex optionally comprises a complete or partial Signal Domain. The term "a Complex comprising partially hybridized pair of nucleic acids" indicates that the nucleic acids of the complex are hybridized to each other over at least five contiguous nucleotides.

**Homopolymeric:** A homopolymeric nucleic acid or nucleic acid domain is a nucleic acid or nucleic acid domain in which the base composition of each of the nucleotides is the same, e.g., poly-A. When a nucleic acid or nucleic acid sequence is said to be x% homopolymeric, then the base composition of x% of the nucleotides of the nucleic acid is the same. Such nucleotides can be but are not necessarily contiguous.

**Target:** As used herein, and unless indicated otherwise, the term "Target" refers to a molecule to which a Probe produced by the methods of the invention is designed to bind. The Target can be any molecule to which a Probe can bind. The Target can be a nucleic acid (a Target Nucleic Acid). For example, the Probe can be utilized for detection of the Target Nucleic Acid in a sample via hybridization to the Probe. The Target can also, for example, be a protein (a Target Protein). The Probe can be utilized, for example, in an electrophoretic mobility shift assay to detect a Target Protein that binds the Probe.

**Probe:** As used herein, the term "Probe" refers to the labeled nucleic acid generated by extension of the Substrate Nucleic Acid according to the methods of the present invention. The term is also meant to encompass a Complex comprising a labeled (i.e., Probe) nucleic acid strand.

**Binding domain:** As used herein, a Binding Domain is a region of complementarity between the Substrate Nucleic Acid and the Template Nucleic Acid. The Binding Domain of the Substrate Nucleic Acid is the Template Binding Domain and the Binding Domain of the Template Nucleic Acid is the Substrate Binding Domain. A Binding Domain can be utilized as part of any Template or Substrate Nucleic Acid molecule, and, therefore, can be considered a "universal" sequence. As such, this term can also be referred to as a "Universal Binding Domain." The term "Template Binding Domain" is used interchangeably with the term "Template Hybridization Domain", and the term "Substrate Binding Domain" is used interchangeably with the term "Substrate Hybridization Domain."

**Adenosine/cytidine/guanosine/thymidine/uridine:** As used herein, adenosine, cytidine, guanosine, thymidine and uridine refer to both the ribonucleosides and the deoxyribonucleosides, unless otherwise indicated.

**Hybridization:** The term "hybridizable to", when referring to the association between the Substrate and Template Nucleic Acids, indicates hybridization under

conditions of moderate to high stringency. Conditions of moderate to high stringency which may be used depend on the nature of the nucleic acid (*e.g.* length, GC content, *etc.*) and are well known in the art. For example, stringent hybridization of a Template Nucleic Acid to a Substrate Nucleic Acid over a region of approximately 15 bases is done under the following conditions: a salt concentration of 50 mM KCl, a buffer concentration of 10 mM Tris-HCl, a  $Mg^{2+}$  concentration of 1.5 mM, a pH of 7-7.5 and an annealing temperature of 55-60°C. Selection of other appropriate conditions for such stringencies is well known in the art (*see e.g.*, Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, 2d Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York; *see also*, Ausubel et al., eds., in the Current Protocols in Molecular Biology series of laboratory technique manuals, 1987-present, Current Protocols, John Wiley and Sons, Inc.).

Complementary/Complementarity: These terms, as used herein, refer to a degree of base-pairing complementarity between two nucleic acid sequences that allows hybridization under the conditions described above. In a preferred embodiment, the degree of complementarity between two nucleic acids is 100%.

#### 4. BRIEF DESCRIPTION OF THE FIGURES

FIG. 1A-D: A schematic representation of a preferred embodiment of the present invention. (A(i)) The design of the Substrate Nucleic Acid where the Ns can be any nucleotide and represent the Target Binding Domain of the Substrate Nucleic Acid and the hatched box represents the Universal Template Hybridization Domain. (A(ii)) The Template Nucleic Acid, where the hatched box represents a sequence complementary to the Template Hybridization Domain in (1A) (the Substrate Hybridization Domain) and the 5' oligo-dT represents the Signal Template Domain. (B) Annealing of the Substrate and Template Nucleic Acids at their respective binding sites. (C) Labeling of the Substrate by primer extension on the oligo-dT Template in the presence of buffer,  $\alpha$ - $^{32}P$ -dATP, and DNA polymerase. (D) Removal of unincorporated  $\alpha$ - $^{32}P$ -dATP by column chromatography and, optionally, removal of the Template Nucleic Acid and any unlabeled products by preparative PAGE to yield a highly Probe with a specific activity many fold higher than when labeled by conventional methods.

FIG. 2A-B: Evaluation of primer extension reaction products. (A) Substrate and Template Nucleic Acid Strands. Template Nucleic Acids in which Substrate Hybridization Domains were of 4, 5, or 6 nucleotides in length were compared for efficiency in promoting primer extension of the Substrate Nucleic Acid. The nucleotides of the binding domain within the Template Nucleic Acid strand were modified to include deoxycytidine (dC), C5-methyl-dC, (5-me-dC) or C5-propynyl-dC as indicated. In this study, the Substrate Nucleic

Acid was initially labeled at the 5' end with  $^{32}\text{PO}_4$  by PNK and then extended with unlabeled dATP in order to define relative reaction efficiencies of Template Nucleic Acids which varied in the length and chemical composition of their Substrate Hybridization Domains.

(B) PAGE analysis of reaction products. Reaction products were separated by analytical  
5 PAGE and visualized by phosphorimaging.

FIG. 3: The percent of the total Substrate Nucleic Acid that is extended in the primer extension (labeling) reaction. Bands from the image shown in FIG. 2 were quantitatively measured using phosphorimaging and a graphical depiction is shown as the percent ratio of Substrate Nucleic Acid extended to total Substrate Nucleic Acid in the  
10 reaction mixture.

FIG. 4A-B: Dot blot assay comparing the detection efficiency (sensitivity) of a Target Nucleic Acid by a Probe labeled by the methods of the present invention compared to a probe labeled with T4 PNK. A) Probes: A Substrate oligonucleotide with a Target Binding Domain complementary to the plasmid target (pGreen Lantern<sup>TM</sup>-1, Life  
15 Technologies, Inc.) was labeled using the methods described herein or by T4 PNK. The base composition and radioactive  $^{32}\text{P}$  label content of both probes are shown. B) Dot Blot Assay: A decreasing mass-titration of the plasmid target was spotted onto two nylon membranes in amounts as indicated. An equal amount of radioactivity per probe (as measured by counts per minute (CPM)), was hybridized to the blots. Both blots were pre-  
20 hybridized, hybridized, and washed in an identical fashion as described in Example 2. Results were visualized using a Packard Cyclone Phosphor Imaging system.

FIG. 5A-B: Northern blot of human placental RNA comparing the detection efficiency (sensitivity) of a Target Nucleic Acid by a Probe labeled by the methods of the present invention with T4 PNK. (A) Probes: A Substrate oligonucleotide with a Target  
25 Binding Domain complementary to human  $\beta$ -actin was labeled using the methods described herein or using T4 PNK. The base composition and radioactive  $^{32}\text{P}$  label content of both probes are shown. (B) Northern Blot: Ten micrograms of human placental total RNA was separated on each lane of a formaldehyde agarose gel and transferred to nylon membranes. An equal amount radioactivity per probe (as measured by counts per minute (CPM)) was  
30 hybridized to the blots. Both blots were pre-hybridized, hybridized, and washed in an identical fashion as described in Example 3. Results were visualized using a Packard Cyclone Phosphor Imaging system. Position of 18S and 28S ribosomal RNA bands are indicated as size markers.

FIG. 6A-B: Primer extension reaction using an RNA Substrate Nucleic Acid. (A)  
35 Substrate Nucleic Acids and Reaction Products. DNA and RNA Substrate Nucleic Acid molecules were initially labeled at the 5' end with  $^{32}\text{PO}_4$  by PNK and then extended with

unlabeled dATP in order to define an RNA substrate composition. (B) PAGE analysis of reaction products. Reaction products were separated by analytical PAGE and visualized by phosphorimaging. Position of Substrate and Probe bands are indicated.

5           **5.     DETAILED DESCRIPTION OF THE INVENTION**

The present invention is directed to methods for the generation of nucleic acid probes. The invention is also directed to nucleic acid probes generated by the methods of the invention, reaction mixtures for producing the probes, and kits comprising reagents for the generation of such probes.

10           The present invention provides a method of labeling a nucleic acid molecule, comprising the steps of hybridizing a first nucleic acid, the Template Nucleic Acid, to a second nucleic acid, the Substrate Nucleic Acid. The Template Nucleic Acid comprises, from 3' to 5', a Substrate Hybridization Domain and a Signal Template Domain, wherein the Substrate Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides  
15 and the Signal Template Domain comprises a sequence of about 5 to about 100 nucleotides. The Substrate Nucleic Acid comprises, from 3' to 5', a Template Hybridization Domain and a Target binding domain, wherein the Template Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides, is not detectably labeled, and shows complementarity toward and is hybridizable to the Substrate Hybridization Domain of the  
20 Template Nucleic Acid, and the Target Binding Domain is not detectably labeled and comprises a nucleotide sequence heterologous to that of the Template Hybridization Domain of the Substrate Nucleic Acid. After the Template Nucleic Acid is hybridized to the Substrate Nucleic Acid, the Substrate Nucleic Acid is extended in the presence of at least one labeled nucleotide triphosphate to create a Signal Domain having a sequence  
25 which shows complementarity toward and is hybridizable to the Signal Template Domain of the Template Nucleic Acid. The resulting Probe can be purified from any remaining nucleic acid for example by PAGE, to obtain a known species of high specific activity. Similarly, the Probe can be isolated from the Complex (*i.e.*, the Probe-Template Nucleic Acid Complex) produced by the extension reaction, *e.g.*, via denaturation (for example by  
30 heat or alkali) followed by such Probe strand purification.

The present invention further provides a Probe comprising, from 3' to 5', a Signal Domain, a Template Hybridization Domain and a Target Binding Domain, wherein the Signal Domain comprises a sequence of about 5 to about 100 nucleotides of which at least two nucleotides are detectably labeled; the Template Hybridization Domain comprises a  
35 sequence of about 5 to about 20 nucleotides, and is not detectably labeled; and the Target



Binding Domain is not detectably labeled and comprises a nucleotide sequence heterologous to that of the Template Hybridization Domain.

In one embodiment, the nucleotides which comprise the Probe are deoxyribonucleotides. In another embodiment, the nucleotides which comprise the Signal Domain of the Probe are deoxyribonucleotides and the nucleotides which comprise the Template Hybridization Domain and the Target Binding Domain of the Probe are ribonucleotides. In certain embodiments of the invention, the Signal Domain of the Probe comprises a sequence of about 10 to about 50 nucleotides.

The Signal Domain of the invention comprises at least two labeled nucleotides. In certain embodiments, the Signal Domain comprises at least three labeled nucleotides, more preferably at least four labeled nucleotides, more preferably at least five labeled nucleotides, and yet more preferably about 10 labeled nucleotides. In other embodiments, the Signal Domain comprises fewer than 20 labeled nucleotides. In yet other embodiments, all the nucleotides which comprise the Signal Domain of the invention are labeled.

The present invention provides a Probe in which the Signal Domain is at least 50%, at least 70%, at least 90% or 100% homopolymeric.

The present invention provides a Probe in which at least one nucleotide comprises a label selected from the group consisting of:  $^{32}\text{P}$ ,  $^{33}\text{P}$ ,  $^{35}\text{S}$ , fluorescein, digoxigenin, biotin, Cy5, Cy3, and rhodamine.

The present invention provides a Probe in which the Template Hybridization Domain comprises a sequence of about 5 to about 10 nucleotides.

The present invention further provides a Probe in which at least 60%, at least 80% or 100% of the nucleotides of the Template Hybridization Domain comprise guanosine or cytidine or a combination thereof. In instances wherein 100% of the nucleotides of the Template Hybridization Domain comprise guanosine or cytidine, it is preferred that there be a combination of guanosine and cytidine present.

In certain embodiments of the present invention, at least 60% of the nucleotides of the Template Hybridization Domain of the Probe comprise guanosine or cytidine or a combination thereof, and the Signal Domain of the Probe is at least 50% homopolymeric.

The present invention further provides a Complex, comprising a Template Nucleic Acid comprising, from 3' to 5' a Substrate Hybridization Domain and a Signal Template Domain, wherein the Substrate Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides and the Signal Template Domain comprises a sequence of about 5 to about 100 nucleotides; and a Probe comprising from 3' to 5' a Signal Domain, a Template Hybridization Domain and a Target Binding Domain, wherein the Signal Domain comprises a sequence of about 5 to about 100 nucleotides, which sequence shows

complementarity toward and is hybridizable to the Signal Template Domain of the Template Nucleic Acid, and of which at least two nucleotides are detectably labeled; the Template Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides, is not detectably labeled, and shows complementarity toward and is hybridizable to the

5 Substrate Hybridization Domain of the Template Nucleic Acid; the Target Binding Domain is not detectably labeled and comprises a nucleotide sequence heterologous to that of the Template Hybridization Domain of the Substrate Nucleic Acid.

The present invention provides a Complex in which the nucleotides which comprise the Probe or Template Nucleic Acid are deoxyribonucleotides. The present invention

10 provides a Complex in which the nucleotides which comprise Template Nucleic Acid are ribonucleotides. The present invention provides a Complex in which the nucleotides which comprise the Signal Domain of the Probe are deoxyribonucleotides and the nucleotides which comprise the Template Hybridization Domain and the Target Binding Domain of the Probe are ribonucleotides.

15 The present invention provides a Complex in which the Substrate Hybridization Domain is at the 3' end of the Template Nucleic Acid.

The present invention provides a Complex in which the Substrate Hybridization Domain of the Template Nucleic Acid comprises a sequence of about 5 to about 10 nucleotides.

20 The present invention provides a Complex in which the Substrate Hybridization Domain of the Template Nucleic Acid cannot be extended by a 5'-3' DNA polymerase.

The present invention provides a Complex in which the Substrate Hybridization Domain of the Template Nucleic Acid further comprises an extension of nucleotides at the 3' end of said Substrate Hybridization Domain, the extension having no complementarity to

25 the Template Hybridization Domain of the Substrate Nucleic Acid.

the Substrate Hybridization Domain comprises a 3'-terminal modified nucleotide. The modification is selected from the group consisting of: a 3'-amino-modifier, a 2', 3'-dideoxynucleotide, a 3'-phosphate, and a modified 3'-phosphate group.

The present invention provides a Complex in which the Substrate Hybridization

30 Domain comprises at least one nucleotide which comprises a modified cytidine, which nucleotide is selected from the group consisting of: C5-methyl-dC and C5-propynyl-dC.

The present invention further provides a Complex in which the Signal Domain of the Probe comprises a sequence of about 10 to about 50 nucleotides.

The present invention further provides a Complex in which the Signal Domain is at

35 least 50%, at least 70%, at least 90% or 100% homopolymeric.

The present invention further provides a Complex in which the at least 60%, at least 80% or 100% of the nucleotides of the Template Hybridization Domain comprise guanosine or cytidine or a combination thereof. In instances wherein 100% of the nucleotides of the Template Hybridization Domain comprise guanosine or cytidine, it is preferred that there be  
5 a combination of guanosine and cytidine present.

The present invention yet further provides a Complex in which the at least 60% of the nucleotides of the Template Hybridization Domain comprise guanosine or cytidine or a combination thereof, and the Signal Domain is at least 50% homopolymeric.

The present invention further provides a reaction mixture for use in a process for the  
10 labeling of a nucleic acid molecule comprising a Template Nucleic Acid comprising, from 3' to 5', a Substrate Hybridization Domain and a Signal Template Domain, wherein the Substrate Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides, and the Signal Template Domain comprises a sequence of about 5 to about 100 nucleotides; and a Substrate Nucleic Acid comprising from 3' to 5', a Template Hybridization Domain  
15 and a Target Binding Domain, wherein the Template Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides, is not detectably labeled, and shows complementarity toward and is hybridizable to the Substrate Hybridization Domain of the Template Nucleic Acid; the Target Binding Domain is not detectably labeled and comprises a nucleotide sequence heterologous to that of the Template Hybridization Domain of the  
20 Substrate Nucleic Acid.

The present invention further provides a method for detecting a Target Nucleic Acid in a sample, comprising contacting the sample with a Probe (comprising a Substrate Nucleic Acid comprising a Signal Domain, which is optionally hybridized or cross linked to a Template Nucleic Acid) under conditions whereby said Probe can bind to the Target  
25 Nucleic Acid to form a Probe-Target hybrid; and detecting any Probe-Target Nucleic Acid hybrids, so that if a Probe-Target Nucleic Acid hybrid is detected, a Target Nucleic Acid is detected in the sample.

The present invention further provides a kit for labeling a nucleic acid molecule, comprising a reaction mixture comprising Template Nucleic Acid comprising, from 3' to 5',  
30 a Substrate Hybridization Domain and a Signal Template Domain, wherein the Substrate Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides, and the Signal Template Domain comprises a sequence of about 5 to about 100 nucleotides; a Substrate Nucleic Acid comprising from 3' to 5', a Template Hybridization Domain and a Target Binding Domain, wherein the Template Hybridization Domain comprises a sequence  
35 of about 5 to about 20 nucleotides, is not detectably labeled, and shows complementarity toward and is hybridizable to the Substrate Hybridization Domain of the Template Nucleic

Acid; the Target Binding Domain is not detectably labeled and comprises a nucleotide sequence heterologous to that of the Template Hybridization Domain of the Substrate Nucleic Acid; and a DNA polymerase.

5                   **5.1    SUBSTRATE NUCLEIC ACID AND TEMPLATE NUCLEIC  
                  ACID SEQUENCES**

          The present invention provides a Template Nucleic Acid comprising, from 3' to 5', a Substrate Hybridization Domain and a Signal Template Domain, wherein the Substrate Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides and the  
10   Signal Template Domain comprises a sequence of about 5 to about 100 nucleotides. The present invention also provides a Substrate Nucleic Acid comprising, from 3' to 5', a Template Hybridization Domain and a Target Binding Domain, wherein the Template Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides, is not  
15   detectably labeled, and shows complementarity toward and is hybridizable to the Substrate Hybridization Domain of the Template Nucleic Acid, and the Target Binding Domain is not detectably labeled and comprises a nucleotide sequence heterologous to that of the Template Hybridization Domain of the Substrate Nucleic Acid.

          The Substrate and Template Nucleic Acids of the invention can comprise ribonucleotides or deoxyribonucleotides, or a combination thereof. The Probes of the  
20   invention can comprise deoxyribonucleotides or a combination of deoxyribonucleotides and ribonucleotides, wherein the nucleotides which comprise the Signal Domain are deoxyribonucleotides. In one embodiment, the nucleotides which comprise the Signal Domain of a Probe are deoxyribonucleotides and the nucleotides which comprise the Template Hybridization Domain and the Target Binding Domain of the same Probe are  
25   ribonucleotides.

          The Substrate Nucleic Acid of the invention can be from about 15 to about 10,000 nucleotides. In certain embodiments, the Substrate Nucleic Acid is about 15 to about 150 nucleotides. In other embodiments, the Substrate Nucleic Acid is about 50 to about 10,000 nucleotides. Preferably, the Substrate Nucleic Acid is about 15 to about 500 nucleotides,  
30   more preferably about 20 to about 300 nucleotides, more preferably from about 25 to about 200 nucleotides, and most preferably from about 30 to about 100 nucleotides.

          The Target Binding Domain of the invention can be from about 10 to about 10,000 nucleotides. In certain embodiments, the Target Binding Domain is about 10 to about 150 nucleotides. In other embodiments, the Target Binding Domain is about 50 to about 10,000  
35   nucleotides. Preferably, the Target Binding Domain is about 10 to about 500 nucleotides, more preferably about 15 to about 300 nucleotides, more preferably from about 20 to about 200 nucleotides, and most preferably from about 25 to about 100 nucleotides.

### 5.1.1 **TEMPLATE AND SUBSTRATE HYBRIDIZATION DOMAINS**

The Substrate Hybridization and Template Hybridization Domain sequences are preferably short so as not to interfere with the use of the Template Nucleic Acid/Probe Complex as hybridization probe. A Substrate or Template Hybridization Domain of about 5 to about 20 bases in length is preferred, with a length of about 5 to about 10 bases in length being most preferred. Sequences shorter than 5 bases can be used but generally give lower yields of the labeled extension product (FIG. 2 and FIG. 3). Sequences longer than 10 bases can also be used, but are preferably no longer than approximately 20 nucleotides.

Preferably, at least 60%, more preferably at least 80% , and most preferably 100% of the nucleotides of the Template Hybridization Domain comprise guanosine or cytidine or a combination thereof. In instances wherein 100% of the nucleotides of the Template Hybridization Domain comprise guanosine or cytidine, it is preferred that there be a combination of guanosine and cytidine present.

In one embodiment of the invention, at least 60% of the nucleotides of the Template Hybridization Domain comprise guanosine or cytidine or a combination thereof, and the Signal Domain is at least 50% homopolymeric.

In a preferred embodiment, the Substrate Hybridization Domain is located at the 3' of the Template Nucleic Acid. In other embodiments, the Substrate Hybridization Domain is located near but not at the 3' of the Template Nucleic Acid. In the latter embodiment a 3' overhang results when the Template Nucleic Acid is hybridized to the Substrate Nucleic Acid. A 3' overhang that is not complementary to the Target Hybridization Domain of the Substrate Nucleic Acid can serve to functionally block the 3' end of the Template Nucleic Acid from primer extension by the polymerase used in the extension reaction.

In certain embodiments of the invention, the Substrate Hybridization Domain cannot be extended by a 5'→3' DNA. In one embodiment, the Substrate Hybridization Domain cannot be extended by a 5'→3' DNA polymerase because it further comprises an extension of nucleotides at the 3' end of the Substrate Hybridization Domain, the extension having no complementarity to the Template Hybridization Domain of the Substrate Nucleic Acid, as discussed above. The 3' extension of the Substrate Hybridization Domain is preferably less than 20 nucleotides, more preferably less than 10 nucleotides, more preferably less than 5 nucleotides, and most preferably 2 or 3 nucleotides. In one embodiment, the Substrate Hybridization Domain cannot be extended by a 5'→3' DNA polymerase because the Substrate Hybridization Domain comprises a 3'-terminal modified nucleotide. In certain modes of the embodiment, the modification is selected from the group consisting of a 3'-amino-modifier, a 2', 3'-dideoxynucleotide, a 3'-phosphate, and a modified 3'-phosphate

group. A 3' amino modifier is an aliphatic chain comprising a primary amino group, which amino group is covalently linked to the 3' OH group of a nucleic acid.

In a preferred embodiment, the Substrate Hybridization Domain to be used as a Substrate Hybridization Domain is GGCGGG, (SEQ ID NO:1). Such a domain is entirely  
5 GC in composition, allowing for high affinity binding to the Template Hybridization Domain, which in turn permits the sequence to be minimal in length. Such a domain can only hybridize to the Template Hybridization Domain in a single alignment, and thus no "slippage" can occur. One hundred percent homopolymeric Template Hybridization Domains (e.g., GGGGGG (SEQ ID NO:2) or CCCCCC (SEQ ID NO:3), therefore, are not  
10 preferred. Using the exemplary, non-limiting Substrate and Template Binding Domains (see Section 5.1.3 below), all Template Nucleic Acid - Substrate Nucleic Acid Complexes formed are optimally aligned and can, therefore, optimally participate in the extension reaction. An additional feature of the preferred Substrate Hybridization Domain is that it is guanosine rich, and thus its complement in the Template Nucleic Acid is cytidine rich. This  
15 allows for incorporation of dC-heterocyclic base-modifications into the Template Nucleic Acid as opposed to the Substrate Nucleic Acid. Use of costly modified-base reagents, such as C5-propynyl-dC or C5-methyl-dC, is more economical when incorporated into the Template Nucleic Acid rather than the Substrate Nucleic Acid, since, because any Substrate Nucleic Acid can be synthesized to include a 3' Template Hybridization Domain, a single  
20 heterologous Template can be used in hundreds or thousands of different Substrate Nucleic Acids and can be utilized for a wide range of labeling reactions, including labeling reactions that simultaneously or individually generate different Probes.

The 3' end of the Template Nucleic Acid can be blocked to prevent the Substrate Nucleic Acid from acting as a template in an undesired primer extension reaction and  
25 thereby labeling the Template Nucleic Acid as well. Function blocking can be achieved by the addition of non-complementary bases at the 3' end as described above. Blocking can also be achieved through chemical modification of the 3' terminal sugar group. Blocking groups which may be used include, but are not limited to, 3'-amino-modifiers, a 2',3'-dideoxynucleotide, a 3'-phosphate or a modified 3'-phosphate group (i.e., -P-O-R, where R  
30 is selected from an alkyl group, an alcohol (e.g. CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>OH) or a derivative of ethylene glycol, (e.g., CH<sub>2</sub>CH<sub>2</sub>-O-). While a blocking group is not required its use eliminates the potential for undesired side reactions.

In certain preferred embodiments of the invention, the Substrate Hybridization Domain comprises at least one nucleotide which comprises a modified cytidine, which  
35 nucleotide is selected from the group consisting of: C5-methyl-dC and C5-propynyl-dC.

### 5.1.2 SIGNAL TEMPLATE AND SIGNAL DOMAINS

According to the present invention, the Signal Domain of the Probe comprises a sequence of about 5 to about 100 nucleotides, which sequence shows complementarity toward and is hybridizable to the Signal Template Domain of the Template Nucleic Acid, of which at least two nucleotides are detectably labeled. Likewise, the size of the Signal Template Domain is generally from about 5 to about 100 nucleotides.

In a preferred embodiment, the Signal Template or Signal Domain is at least 50% homopolymeric, more preferably at least 70% homopolymeric, more preferably at least 90% homopolymeric, and most preferably is a homopolymer (*e.g.*, poly-A). In instances wherein a Probe of the invention is to be utilized for detecting Targets in a sample that contain or are suspected of containing mRNA species, a homopolymeric poly-T or poly-dT Signal Domain is not preferred. An advantage of a homopolymeric Signal Domain is that only one type of nucleotide needs to be used in the extension reaction that generates the probe, for example,  $^{32}\text{P}$ -dATP. In a preferred embodiment, the Signal Template Domain is oligo-dT (or oligo-U).

The Signal Template domain or the Signal Domain is preferably from about 5 to about 100 nucleotides, more preferably from about 5 to about 75 nucleotides, more preferably from about 5 to about 50 nucleotides, and most preferably from about 5 to about 20 nucleotides.

In one embodiment of the invention, the Substrate Nucleic Acid is extended to produce the Signal Domain using a nucleotide triphosphate that comprises a label selected from the group consisting of  $^{32}\text{P}$ ,  $^{33}\text{P}$ ,  $^{35}\text{S}$ , fluorescein, digoxigenin, biotin, Cy5, Cy3, and rhodamine. In certain modes of the embodiment, the nucleotide triphosphate is selected from the group consisting of  $\alpha$ - $^{32}\text{P}$ -dATP,  $\alpha$ - $^{33}\text{P}$ -dATP,  $\alpha$ - $^{35}\text{S}$ -dATP, fluorescein-15-dATP, digoxigenin-16-dATP, biotin-7-dATP, and biotin-21-dATP.

### 5.1.3 EXEMPLARY EMBODIMENT

In an exemplary embodiment, presented herein by way of example and not by way of limitation, the Substrate and Template Nucleic Acid Comprise the following sequences (in which N is any ribonucleotide or deoxyribonucleotide):

#### Target Binding Domain of Substrate Nucleic Acid

5'-NNNNNNNNNNNNNNNNNNNNNNNNNNNN-3' (SEQ ID NO:4)

#### Template Hybridization Domain of Substrate Nucleic Acid:

5'-GGCGGG-3' (SEQ ID NO:1)

#### Substrate Nucleic Acid:

5'-NNNNNNNNNNNNNNNNNNNNNNNNNGGCGGG3' (SEQ ID NO:5)

Signal Template Domain of Template Nucleic Acid:

5'-TTTTTTTTTT-3' (SEQ ID NO:6)

Substrate Hybridization Domain of Template Nucleic Acid:5'-CCCGCC-C<sub>7</sub>-amino-modifier-3' (SEQ ID NO:7)5 Template Nucleic Acid:5'-TTTTTTTTTTCCCGCC-C<sub>7</sub>-amino-modifier-3' (SEQ ID NO:8)

In a preferred embodiment the Template Nucleic Acid is modified to improve labeling efficiency. In a preferred mode of the embodiment, the modification is the substitution of the dC nucleotide of the Substrate Binding Domain with 5-methyl-deoxycytidine (mC), or even more preferably, with C5-propynyl-dC (pC). In the exemplary embodiment above, modification of the Template Nucleic Acid with mC or pC would result in:

Template Nucleic Acid with methyl-dC modifier:5'-TTTTTTTTTTmCmCmCGmCmC-C<sub>7</sub>-amino-mod-3' (SEQ ID NO: 9)15 Template Nucleic Acid (propynyl-dC modifier):5'-TTTTTTTTTTpCpCpCGpCpC-C<sub>7</sub>-amino-mod-3' (SEQ ID NO: 10)

These heterocyclic base modifications increase the binding affinity of the Substrate Hybridization Domain for the Template Hybridization Domain, allowing for more complete hybridization and therefore a more efficient labeling reaction. Other heterocyclic modifications can be used, *e.g.* C5-propynyl-dU (as a dT or U substitute), or 2,6-diaminopurine (as a dA or A substitute). Such modifications serve to increase the relative reaction efficiency but are not required for practice of the invention (FIG. 2 and FIG. 3).

The labeling reaction utilizes the synthesis of RNA or DNA to form a covalent attachment between a labeled nucleotide and a substrate nucleic acid by primer extension.

25 The Substrate and Template Nucleic Acids associate with each other, generally through standard Watson-Crick base pairing of 5 or more-base complementary hybridization sequences. In certain embodiments, the Substrate and Template Nucleic Acids associate by non-Watson-Crick hybridization (such as the incorporation of protein nucleic acid residues in the Template Nucleic Acid).

30

**5.2 NUCLEIC ACID SYNTHESIS**

The nucleic acids of the invention may be synthetic oligonucleotides comprised of deoxyribonucleotides, ribonucleotides, combinations (chimeras), or derivatives thereof.

The use of ribonucleotide derivatives has been reported by Shibahara *et al.*, Nucleic Acids Research 1989, 17, 239-252. A person skilled in the art is capable of synthesizing



oligonucleotides. See for example, Gait, M. J., ed. (1984) *Oligonucleotide Synthesis* (IRL, Oxford).

In one preferred embodiment, nucleic acid synthesis is performed using solid phase phosphoramidite monomer chemistry with automated synthesizers. Currently nucleic acid strands up to about 150 nucleotides in length can be synthesized chemically using the phosphoramidite method. In other embodiments, other methods of nucleic acid synthesis are used, such as H-phosphonate chemistry. Chemical synthesis of nucleic acids allows for the production of various forms of the nucleic acids with modified linkages and nonstandard bases attached in chosen places throughout the nucleic acid's entire length (Current Protocols in Molecular Biology, John Wiley & Sons, Inc., incorporated herein by reference).

In certain embodiments, the Substrate and Template Nucleic Acids are not chemically but enzymatically synthesized, for example by a DNA or RNA polymerase. Such Substrate Nucleic Acids may be single or double stranded. In certain modes of this embodiment, the Substrate Nucleic Acid is from about 50 nucleotides to about 10,000 nucleotides, more preferably from about 50 to about 1,000 nucleotides, and most preferably from about 50 to about 200 nucleotides. In a preferred mode of this embodiment, the Substrate Nucleic Acid is generated by polymerase chain reaction (PCR), using standard techniques. The Template Hybridization Domain can be added, for example, by ligation or by incorporating the Template Hybridization Domain sequence into the PCR Primers. If the Substrate Nucleic Acid is a PCR product, it is heated to denature it from its complementary strand to allow it to anneal with the Template Nucleic Acid.

When the Substrate Nucleic Acid or Template Nucleic Acid is a ribonucleic acid, it can be generated in a transcriptional reaction. A synthetic or PCR amplified DNA clone encoding the sequence of a Substrate Nucleic Acid or the Template Nucleic Acid is prepared, said DNA having an initiation site for an RNA polymerase. RNA polymerase initiation sites are well known to those of skill in the art (see *e.g.* Chamberlin and Ryan, 1982, *Bacteriophage DNA-dependent RNA polymerases. In The Enzymes*, Vol. 15B (P.D. Boyer, ed.), pp. 87-109. Academic Press, San Diego. Production of the RNA from the synthetic clone by transcription with the appropriate RNA polymerase (*e.g.* SP6 or T7 RNA polymerase) is achieved using routine approaches which are well known in the art. Suitable RNA polymerases include but are not limited to T3, T7, SP6, BA14, or K11 RNA polymerase. Since RNA polymerases generate heterogeneous 3' ends, it is preferred that the Substrate Nucleic Acid or Template Nucleic Acid generated by an RNA polymerase is initially synthesized as longer precursor, then cleaved by a ribozyme or other appropriate enzyme.

The Substrate and Template Nucleic Acids may also be modified at any position with substituents generally known in the art, with the exception of 2' O-modifications of ribonucleic acids.

In one embodiment, the Substrate and Template Nucleic Acids have one or more  
5 modifications of the nucleotides which increase bioavailability by enhancing stability to  
nuclease attack and/or increasing cellular uptake. For example, the modification of the  
nucleotide backbone bonds from phosphodiester to phosphorothioates confers greater  
stability against nuclease attack. Phosphorothioates, *i.e.*, the substitution of a sulfur atom for  
a phosphate oxygen in the internucleotide phosphodiester linkage, are stable to nuclease  
10 cleavage and soluble in lipid. Lipid solubility can be important for bioavailability.  
Preferably, only the 5' terminal bond of the Substrate and both terminal bonds of the  
Template Nucleic Acids are modified. In this manner, stability is increased without  
foresaking cellular uptake. The oligonucleotides can also be protected from exonuclease  
attack through the addition of amino groups at the ends.

15 Other modified Substrate and Template Nucleic Acids within the scope of the  
invention include  $\alpha$ -anomeric oligonucleotides. An  $\alpha$ -anomeric oligonucleotide forms  
specific double-stranded hybrids with complementary RNA in which, contrary to the usual  
 $\beta$ -units, the strands run parallel to each other (Gautier *et al.*, 1987, Nucl. Acids Res.  
15:6625-6641).

20 Additionally, the Substrate and Template Nucleic Acids may comprise at least one  
modified base moiety which is selected from the group including but not limited to  
5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine,  
4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-  
2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-  
25 galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine,  
2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine,  
5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil,  
5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine,  
5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine,  
30 uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine,  
5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid  
methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-  
carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

The Substrate and Template Nucleic Acids can also comprise at least one  
35 modified sugar moiety selected from the group including but not limited to arabinose,  
2-fluoroarabinose, xylulose, and hexose.

The Template Hybridization Domain, which, acts as a binding site for the Template Nucleic Acid, is preferably added to a nucleic acid at the time the nucleic acid is prepared, for example during chemical synthesis. In other embodiments, the Template Hybridization Domain is added by fill-in reactions, ligation or PCR.

5

### 5.3 PROBE LABELING AND PURIFICATION

Specifically, the invention provides a method of labeling a nucleic acid molecule, comprising the steps of hybridizing a first nucleic acid, the Template Nucleic Acid, to a second nucleic acid, the Substrate Nucleic Acid. The Template Nucleic Acid comprises, from 3' to 5', a Substrate Hybridization Domain and a Signal Template Domain, wherein the Substrate Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides and the Signal Template Domain comprises a sequence of about 5 to about 100 nucleotides. The Substrate Nucleic Acid comprises, from 3' to 5', a Template Hybridization Domain and a Target binding domain, wherein the Template Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides, is not detectably labeled, and shows complementarity toward and is hybridizable to the Substrate Hybridization Domain of the Template Nucleic Acid, and the Target Binding Domain is not detectably labeled and comprises a nucleotide sequence heterologous to that of the Template Hybridization Domain of the Substrate Nucleic Acid. Upon hybridization, the Substrate Nucleic Acid is extended with a DNA polymerase in the presence of a labeled nucleotide to create a Signal Domain having a sequence which shows complementarity toward and is hybridizable to the Signal Template Domain of the Template Nucleic Acid.

The first step of the labeling methods of the present invention, after synthesis of the Substrate and Template Nucleic Acids, is hybridizing the two strands via the Template and Substrate Hybridization Domains, respectively. Hybridizing does not require special conditions and can be done simply by mixing the two nucleic acids together under standard conditions well known to those of skill in the art, *e.g.*, in a low salt, buffered aqueous solution; heat denaturation and cooling may facilitate hybridization. In certain embodiments, hybridization is preceded by heat denaturation and cooling of the nucleic. In a preferred embodiment the Template Nucleic Acid present in the reaction mixture is in excess over the Substrate Nucleic Acid, preferably a 5 to 50-fold molar excess, most preferably a 5 to 25-fold molar excess. Lower molar ratios will still result in labeling, albeit at lower efficiencies. Increasing Template Nucleic Acid above 50-fold molar excess over the Substrate Nucleic Acid does not improve reaction efficiency.

Preferably, the heating step is done in a water bath or other heating device at 94°C for one minute, although different temperatures and incubation times will also suffice, as

can be determined by one of skill in the art. An extended annealing or hybridization step at controlled temperature is not necessary. The reaction buffer is preferably 10 mM Tris pH 7.5, 5 mM MgCl<sub>2</sub>, and 7.5 mM dithiothreitol (DTT). However, other buffers may be utilized, depending on the DNA polymerase of choice and as reaction conditions permit.

- 5 Appropriate conditions for the DNA polymerase of choice are well known to those of skill in the art. The nucleic acid mixture is transferred from the heat source and allowed to cool to room temperature under ambient conditions.

The Substrate Nucleic Acid is then extended with a DNA polymerase in the presence of a labeled nucleotide (see Section 5.5, below) to create a Signal Domain having a  
10 sequence which shows complementarity toward and is hybridizable to the Signal Template Domain of the Template Oligonucleotide.

The precise conditions of any given extension reaction depend on the particular DNA polymerase employed are known to one of skill in the art.

- In a preferred embodiment, the exonuclease negative (exo<sup>-</sup>) Klenow fragment of  
15 *E. coli* DNA polymerase I (1  $\mu$ l, 5 units) and  $\alpha$ -<sup>32</sup>P-dATP (6  $\mu$ l, 10  $\mu$ Ci/ $\mu$ l, final concentration of 1  $\mu$ M) are added to pre-annealed Substrate and Template Nucleic Acids in a total reaction volume of 10  $\mu$ l (see Table 1 and Section 6.3 below). Preferably, highest available specific activity radio nucleotide is employed (currently 6000 Ci/mmol) to obtain the highest sensitivity probes.

20

**TABLE 1: REACTION COMPONENTS**

Component	Concentration	Vol.	Final Concentration
Exo <sup>-</sup> Klenow DNA Polymerase	5 U/ $\mu$ l	1 $\mu$ l	0.5 U/ $\mu$ l
25 $\alpha$ - <sup>32</sup> P-dATP, 6000 Ci/mmol	10 $\mu$ Ci/ $\mu$ l	6 $\mu$ l	6 $\mu$ Ci/ $\mu$ l, 1 $\mu$ M
Substrate Nucleic Acid	0.5 mole	1 $\mu$ l	50 nM
Template Nucleic Acid	12.5 pmole	1 $\mu$ l	1.25 $\mu$ M
30 Reaction Buffer	10x	1 $\mu$ l	10 mM Tris pH 7.5 5 mM MgCl <sub>2</sub> 7.5 mM DTT
		10 $\mu$ l Final Vol.	

35

The present invention encompasses methods that may yield less than 90% labeling efficiency. When the labeling efficiency is less than 90%, the labeled nucleic acid is

preferably isolated from the unlabeled nucleic acid. The difference between the sizes of the Substrate Nucleic Acid and the Probe allows the probe to be separated from the unlabeled Substrate Nucleic Acid by, for example, preparative polyacrylamide gel electrophoresis (PAGE). The Probe migrates at a slower rate on PAGE and can be identified by, for example, its radioactivity. The desired nucleic acid species are cut out of the gel and eluted by standard methods including, but not limited to, electro elution or crushing then soaking the gel slice in water or buffer.

In another embodiment, the Probe is a Complex that also includes the Template Nucleic Acid. The double stranded portion of the Complex will not interfere with the hybridization of the Target Binding Domain in subsequent assays. Optionally, after the Signal Domain is formed, the Probe is covalently cross-linked to the Template Nucleic Acid. Nucleic acid cross-linking agents suitable for the present invention include, but are not limited to, 4'-aminomethyltrioxsalen hydrochloride (Sigma-Aldrich, Milwaukee, WI), N-5-azido-2-nitrobenzoyloxysuccinimide, dimethyl adipimidate-2HCl, and dimethyl pimelimidate-2HCl (Pierce).

#### 5.4 POLYMERASES FOR THE EXTENSION REACTION

In a preferred embodiment, the enzyme used to catalyze the primer extension labeling reaction (see FIG. 1B and FIG. 1C) is an *exo* Klenow fragment of *E. coli* DNA polymerase I. In alternative embodiments, other DNA polymerase enzymes can be used including viral (bacteriophage) DNA polymerases, thermophilic DNA polymerases (such as those used in PCR), and reverse transcriptase. These enzymes may be native, recombinant, or modified. Modifications of DNA polymerases contemplated by the methods of the present invention include but are not limited to modifications that remove activities other than polymerase activities, for example exonuclease activity. Suitable DNA polymerases for the primer extension reaction include but are not limited to *E. coli* DNA Polymerase I, T4 DNA polymerase, T7 DNA polymerase (native or modified), *Taq* DNA polymerase, or reverse transcriptase. One skilled in the art would be able to identify other polymerases suitable for the primer extension reaction of the present invention.

It is advantageous that the DNA polymerase used in the labeling reaction be a processive enzyme, as are those listed above. This ensures that the primer extension reaction is complete and that only the full-length extension product is formed.

## 5.5 NUCLEOTIDE LABELS AND DETECTION

### 5.5.1 RADIOACTIVITY

In certain embodiments, the labeled nucleotide incorporated into the Signal Domain is a radiolabeled nucleotide. In a preferred embodiment, the labeled nucleotide is  $\alpha$ -<sup>32</sup>P-dATP. However, any other labeling moiety that is adaptable to the present protocol can be used, as will be recognized by one of skill in the art. Commonly used radioactively labeled nucleotides for the practice in the invention include, but are not limited to,  $\alpha$ -<sup>32</sup>P-dNTP,  $\alpha$ -<sup>33</sup>P'-dNTP, and  $\alpha$ -<sup>35</sup>S-dNTP. As an alternative to radioactive labeling, non-radioactively labeled nucleotides can be used. For example, biotin-7-dATP functions interchangeably with  $\alpha$ -<sup>32</sup>P-dATP in the labeling reaction (see Section 7, below).

### 5.5.2 FLUORESCENCE

In certain embodiments of the invention, the extension reaction incorporates a fluorescently labeled nucleotide into the Signal Domain. A fluorescently labeled nucleotide includes but is not limited to fluorescein-15-dATP, fluorescein-12-dUTP, tetramethylrhodamine-6-dUTP (Boehringer Mannheim), or Cy5-dCTP (Pharmacia). In a preferred embodiment, the fluorescently labeled nucleotide is fluorescein-15-dATP.

### 5.5.3 HAPTENS AND ENZYMES

Nonisotopically labeled probes typically have a shelf life of over two years and much more DNA can be labeled in one reaction, providing a constant, quality controlled reagent for use in multiple experiments. Some examples of derivative dATP molecules that are not meant to be limiting to the scope of the present invention include, digoxigenin-16-dATP (Boehringer Mannheim), biotin-7-dATP (Sigma, St. Louis, MO), and biotin-21-dATP (Clontech).

One of the ways in which a nucleic acid or antibody can be detectably labeled is by linking the same to an enzyme and use in an enzyme immunoassay (EIA) (Voller, "The Enzyme Linked Immunosorbent Assay (ELISA)", 1978, Diagnostic Horizons 2:1-7, Microbiological Associates Quarterly Publication, Walkersville, Md.); Voller *et al.*, 1978, J. Clin. Pathol. 31:507-520; Butler, 1981, Meth. Enzymol. 73:482-523; Maggio (ed.), 1980, Enzyme Immunoassay, CRC Press, Boca Raton, Fla.; Ishikawa *et al.*, (eds.), 1981, Enzyme Immunoassay, Kigaku Shoin, Tokyo). The enzyme which is bound to the nucleic acid or antibody will react with an appropriate substrate, preferably a chromogenic substrate, in such a manner as to produce a chemical moiety which can be detected, for example, by spectrophotometric, fluorimetric or by visual means. Enzymes which can be used to detectably label the nucleic acid or antibody include, but are not limited to, malate

dehydrogenase, staphylococcal nuclease, delta-5-steroid isomerase, yeast alcohol dehydrogenase, alpha-glycerophosphate, dehydrogenase, triose phosphate isomerase, horseradish peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase and acetylcholinesterase. The detection can be accomplished by colorimetric methods which employ a chromogenic substrate for the enzyme. Detection may also be accomplished by visual comparison of the extent of enzymatic reaction of a substrate in comparison with similarly prepared standards.

## 5.6 KITS

The present invention also provides a kit for the practice of the labeling methods of the invention. The kit comprises in one or more containers a Template Nucleic Acid (preferably universal labeling reagent) and one or more of the following: (i) 10x stock buffer (e.g. 100 mM Tris pH 7.5; 50 mM MgCl<sub>2</sub>, 75 mM DTT) and a stop solution (10 mM EDTA)); (ii) RNA or DNA polymerase (most preferably the exo<sup>-</sup> Klenow Fragment of *E. coli* DNA polymerase); (iii) a Template Hybridization Domain for ligating to a Target Binding Domain to generate a Substrate Nucleic Acid; (iv) a control Substrate Nucleic Acid and a control Target Nucleic Acid; and/or (v) a labeled nucleotide for use in the extension reaction and, optionally, if the label is a hapten, an antibody for detection, or if the label is an enzyme, colorimetric reagents for detection.

In a preferred embodiment, a kit of the invention comprises a reaction mixture comprising a Substrate Nucleic Acid and Template Nucleic Acid, and a DNA polymerase. Optionally, the kit further comprises a buffer, a stop buffer, a labeled nucleotide, a control substrate and control target nucleic acid, and/or reagents for the detection of a labeled nucleotide.

The following experimental examples are offered by way of illustration and not by way of limitation.

## 6. EXAMPLE 1: EFFICIENCY OF THE PRIMER EXTENSION REACTION WITH SUBSTRATE AND TEMPLATE NUCLEIC ACIDS

## 6.1 MATERIALS AND REAGENTS

The following reagents were used for polynucleotide (PNK) reactions: an oligonucleotide substrate; T4 PNK (New England Biolabs);  $\gamma$ -<sup>32</sup>P-ATP, 6000 Ci/mmol (Amersham); the buffer: 70 mM Tris pH 7.6, 10 mM MgCl<sub>2</sub>, and 5 mM DTT; TE (10 mM Tris pH 8.0, 1 mM EDTA); and, distilled, deionized water.

The following reagents were used for primer extension reactions: a PAGE purified Substrate oligonucleotide with a 3' Binding Domain; PAGE purified Template Nucleic Acid; 10x stock buffer (100 mM Tris pH 7.5; 50 mM Mg Cl<sub>2</sub> 75 mM DTT); a stop solution (10 mM EDTA); 10  $\mu$ M dATP (unlabeled); and exo<sup>-</sup> Klenow DNA Polymerase (5 units/ $\mu$ l, New England Biolabs).

Equipment and supplies included: 500  $\mu$ l microfuge tubes; 1.5 ml screw cap microfuge tubes; Sephadex (G25 spin columns); a waterbath set to 94°C; a waterbath set to 37°C; a tabletop microcentrifuge; and a micropipette with aerosol barrier tips.

Nine distinct Template Nucleic Acids were employed that each contained a 6 (SEQ ID NO:8, 9, 10), 5 (SEQ ID NO:11, 12, 13) or 4 (SEQ ID NO:14, 15, 16) base Substrate Hybridization Domain (FIG. 2A). These Template Nucleic Acids were modified such that the binding domain comprised dC (SEQ ID NO:8, 11, 14), 5-me-dC (SEQ ID NO:9, 12, 15), or C5-propynyl-dC (SEQ ID NO:10, 13, 16). Each Template Nucleic Acid was blocked at the 3' end with an amino modifier-C<sub>7</sub> group (Glen Research).

The Substrate Nucleic Acid was a 36-mer comprised of a unique 30 nucleotide sequence at the 5'-end and a 6 base 3' Template Hybridization Domain (underlined): 5'-CTTGATTAGGGTGATGGTTCACGTAGTGGGGGCGGG-3'. (SEQ ID NO:17)

The purpose of this study was to define the efficiency of the primer extension reaction using the methods of the invention. It was therefore necessary to quantitatively compare unreacted substrate with reacted substrate. The protocol was modified so that radiolabel was incorporated at the 5'-end of the Substrate Nucleic Acid using T4 PNK and  $\gamma$ -<sup>32</sup>P-ATP, and the subsequent primer extension reaction was conducted using unlabeled (cold) dATP, thereby allowing both reacted and unreacted Substrate nucleic acid to be detected.

## 6.2 LABELING NUCLEIC ACIDS USING T4 KINASE

The Substrate Nucleic Acid was resuspended in TE at a final concentration of 0.5  $\mu$ M (0.5 picomoles/ $\mu$ l). From this stock, 2 picomoles (4  $\mu$ l) was 5' end labeled using T4 PNK. The Substrate Nucleic Acid was incubated with 100  $\mu$ Ci  $\gamma$ -<sup>32</sup>P-ATP and 5 units of T4 PNK in a final reaction volume of 50  $\mu$ l with kinase buffer of composition as described



above for 1 hour at 37°C. Unincorporated  $\gamma$ - $^{32}\text{P}$ -ATP was removed by passage through a Sephadex G-25 spin column. Radiolabel incorporation was measured using scintigraphy.

5'  $^{32}\text{P}$ -end-labeled Substrate Nucleic Acid (prepared as described above) was "spiked" into an aliquot of cold Substrate at 10,000 CPM/ $\mu\text{l}$  (approximately 1 femtomole of radiolabeled compound per 0.5 picomole cold oligonucleotide). All reagents were stored at -20°C and kept in an ice bucket at 0°C until thawed for use.

### 6.3 PRIMER EXTENSION REACTION

The primer extension protocol involved two steps: annealing the Substrate and Template Nucleic Acids, and a primer extension reaction. Following the extension reaction, the reaction products were separated using analytical PAGE. The Probe was visualized using phosphorimaging.

1  $\mu\text{l}$  of Substrate Nucleic Acid, 1  $\mu\text{l}$  of 10x stock reaction buffer, and 1  $\mu\text{l}$  of Template Nucleic Acid (resuspended at 12.5  $\mu\text{M}$  concentration in TE (12.5 picomoles per  $\mu\text{l}$ )) were added together in a 500  $\mu\text{l}$  microcentrifuge tube in the molar concentrations shown in Table 2 (see FIG. 1B). The reaction was mixed without vortexing by pipetting up and down. The tube was then placed in a hot water bath at 94°C for 1 minute. The tube was removed from heat and the annealing reaction proceeded in the same tube for 5 minutes, slowly cooling to room temperature. The tube was then centrifuged briefly to return the contents to the bottom of the tube, which ensured that all nucleic acid and buffer was available for the primer extension reaction.

TABLE 2: ANNEALING

Reagent	Amount	Volume
Substrate Nucleic Acid	0.5 pmole (10,000 CPM)	1 $\mu\text{l}$
Template Nucleic Acid	12.5 pmole	1 $\mu\text{l}$
Reaction Buffer	10 x Stock	1 $\mu\text{l}$
Final Volume		3 $\mu\text{l}$

1  $\mu\text{l}$  of 10  $\mu\text{M}$  cold dATP, 5  $\mu\text{l}$  of water, and 1  $\mu\text{l}$  of exo<sup>-</sup> Klenow fragment of *E. coli* DNA polymerase I were added directly into the previously prepared annealing tube and mixed, as in Table 3, using an up and down pipetting action with an aerosol barrier pipette tip. The reaction was allowed to proceed at room temperature for one hour (FIG. C). The reaction was stopped with the addition of an equal volume (10  $\mu\text{l}$ ) of

gel loading buffer (7M urea, 100 mM EDTA), heated at 94°C for 5 minutes, and separated on a 12% polyacrylamide, 7M urea, 0.5 x TBE (45 mM Tris borate, 1 mM EDTA) gel. The gel was exposed to a Packard MP Phosphor Screen for 2 hours and visualized using a Packard Cyclone™ Storage Phosphor System.

5

**TABLE 3: LABELING REACTION**

	Reagent	Volume	Final Concentration
10	Substrate Nucleic Acids, Template Nucleic Acid Buffer	3 $\mu$ l	50 nM Substrate 1.25 $\mu$ M Template 1X Buffer
	Exo <sup>-</sup> Klenow DNA	1 $\mu$ l	0.5 U/ $\mu$ l
	10 $\mu$ M dATP	1 $\mu$ l	1 $\mu$ M
	H <sub>2</sub> O	5 $\mu$ l	
		10 $\mu$ l final vol.	
15			

The Substrate Nucleic Acid, the Template and final reaction products are as follows ("A" = dA residues added by primer extension; "C" = dC, 5-me-dC, or CS-propynyl-dC; "#" = 3'-amino-modifier-C<sub>7</sub>):

**20 Substrate Nucleic Acid (SEQ ID NO.18):**

5' -<sup>32</sup>P-CTTGATTAGGGTGATGGTTCACGTAGTGGGGGCGGG-3'

**Template Nucleic Acid (SEQ ID NOS:8-16):**

3' -#-CCGCCCTTTT TTTT -5'

**Probe (SEO ID NO: 19):**

25 5' -<sup>32</sup>P-CTTGATTAGGGTGATGGTTCACGTAGTGGGGCGGGAAAAAAAAA-3'

3' - # - CCGCCCTTTT TTTT TTTT - 5'

In this example, the original 36-mer Substrate Nucleic Acid was increased in length to 46 bases by the primer extension reaction. The extension efficiency was determined by comparing the amount of nucleic acid signal remaining as a 36-mer (FIG. 2B) to the amount of nucleic acid increased in length to the 46-mer. These values were quantitatively determined by integrating pixel density (ImageQuant™ Software, Packard Instrument Company) from images obtained using a Packard Cyclone™ Storage Phosphor System and are summarized graphically in FIG. 3. Without heterocyclic base modification introduced in the binding domain of the Template Nucleic Acid, a binding domain of 6 bases is necessary to obtain more than 80% extension of the Substrate. The extension efficiency

drops significantly when the binding domain is decreased to 5 nucleic acids and even more so at a length of 4 nucleic acid residues. If 5-me-dC or, even better, CS-propynyl-dC are substituted for deoxycytidine (dC) in the binding domain of the Template Nucleic Acid, the extension efficiency increases. This effect is most evident with a binding domain 5 residues long.

## 7. EXAMPLE 2: COMPARISON OF SENSITIVITY OF PROBES LABELED USING T4 AND PRIMER EXTENSION

### 7.1 MATERIALS AND REAGENTS

Reagents and materials used to label the Substrate Nucleic Acid at the 5' end using T4 PNK and at the 3' end by primer extension were identical to those previously described in Example 1 (Section 6, *supra*). with the exception that  $\alpha$ -<sup>32</sup>P-dATP (6000 Ci/mmol, Amersham) was substituted for unlabeled dATP in the primer extension reaction.

In this example, based on kit format, both the Substrate Nucleic Acid, a 36-mer, and Template Nucleic Acid a 16-mer, are provided lyophilized. The Substrate Nucleic Acid is the same as used in Example 1 and contains a 30 base specific sequence at the 5' end complementary to the target plasmid pGreenLantern™-1 and a 6 base 3' end Template Hybridization Domain, GGCGGG (SEQ ID NO:1). The Template Nucleic Acid consists of a 10 base oligo-dT sequence Signal Template Domain at the 5'-end, a 6 base 3' Substrate Hybridization Domain complementary to the Template Hybridization Domain of the Substrate Nucleic Acid, and is 3'-end blocked with an amino-modifier-C<sub>7</sub> group.

The Substrate Nucleic Acid is:

5' - CTTGATTAGGGTGATGGTTCACGTAGTGGGGGGCGGG 3' (SEQ ID NO:17). The underlined portion is the Template Hybridization Domain. The Template Nucleic Acid is: 5' - TTTTTTTTTTCCCGCC-C<sub>7</sub>-amino-modifier-3' (SEQ ID NO:10), where C represents C5-propynyl-dC.

The Template Nucleic Acid was resuspended in TE at a concentration of 12.5 pM (12.5 picomoles per  $\mu$ l). A stock solution of the Substrate Nucleic Acid was made in TE at a concentration of 10  $\mu$ M (10 picomoles per  $\mu$ l). A working solution of the Substrate was made by 1:20 dilution in distilled water with a final concentration of 0.5  $\mu$ M (0.5 picomoles per  $\mu$ l). All reagents were stored at -20°C and kept in an ice bucket at 0°C until thawed for use. It is preferable to make a series of aliquots of these reagents upon receipt of the kit in single use tubes for both contamination control and ease of use.

### 7.2 LABELING WITH T4 POLYNUCLEOTIDE KINASE

From the 10  $\mu$ M stock of Substrate Nucleic Acid, 2 picomoles (2  $\mu$ l) was 5'-end labeled using T4 PNK, as described in Example 1 (Section 6.2, *supra*). Briefly. Substrate

Nucleic Acid was incubated with 100  $\mu\text{Ci}$   $\gamma$ - $^{32}\text{P}$ -ATP and 5 units of T4 PNK in a final reaction volume of 50  $\mu\text{l}$  with kinase buffer of composition as described for 1 hour at 37°C. Unincorporated  $\gamma$ - $^{32}\text{P}$ -ATP was removed by passage through a Sephadex G-25 spin column. Radiolabel incorporation was measured using scintigraphy and, assuming no loss of labeled nucleic acid or contamination with unincorporated label, resulted in a specific activity of  $1 \times 10^7$  CPM per picomole probe. Structure of this labeled probe is shown in FIG. 4A.

### 7.3 NUCLEIC ACID LABELING BY PRIMER EXTENSION

The labeling protocol involved two steps: annealing the Substrate and Template Nucleic Acid (~10 minutes), and a primer extension reaction (1 hour) and was followed by removal of unincorporated label (~10 minutes).

The annealing step of the primer extension reaction was carried out as described in Example 1 (Section 6.3, *supra*). Briefly, 1  $\mu\text{l}$  (0.5 picomoles) of Substrate Nucleic Acid (from the dilute "working solution"), 1  $\mu\text{l}$  of 10x reaction buffer, and 1  $\mu\text{l}$  (12.5 picomoles) of Template Nucleic Acid were added together in a 500  $\mu\text{l}$  microcentrifuge tube in the same molar concentrations as shown in Table 2. After heating the sample to 94°C (1 minute) and cooling to room temperature (5 minutes), 6  $\mu\text{l}$  of  $\alpha$ - $^{32}\text{P}$  dATP and 1  $\mu\text{l}$  (5 units) of the exo<sup>-</sup> Klenow fragment of *E. coli* DNA polymerase were added. The reaction was allowed to proceed at room temperature for one hour. 40  $\mu\text{l}$  of stop solution (10 mM EDTA) was added and the reaction was passed through a Sephadex G25 spin column to remove unincorporated  $\alpha$ - $^{32}\text{P}$ -dATP. An aliquot was removed and evaluated for radiolabel incorporation by liquid scintillation counting. Assuming no loss of labeled nucleic acid or contamination with unincorporated label, the probe had a specific activity of  $9 \times 10^7$  CPM per picomole, nearly 10x that of the probe labeled with PNK. The structure of the labeled probe produced is shown in FIG. 4A.

### 7.4 COMPARISON OF PROBE SENSITIVITIES IN A DOT BLOT ASSAY

To directly compare the functional sensitivity of a 5' end-labeled probe and a 3' primer-extension labeled probe using the method of the invention, two identical dot blots were prepared using a commercially available plasmid as target, pGreen Lantern<sup>TM</sup>-1. Decreasing mass amounts of plasmid, including 100 attomoles, 10 attomoles, and 1 attomole, were spotted onto MSI nylon transfer membranes (Micron Separations, Inc.), denatured in 0.5 M NaOH, 1.5 M NaCl for 30 minutes, renatured in 1.0 M Tris pH 7.5, 1.5 M NaCl for 30 minutes, and baked at 80°C for 1 hour. The membranes were prehybridized in 0.9 M NaCl, 0.18 M Tris pH 7.5, 0.012 M EDTA, 5x Denhardt's solution, and 100  $\mu\text{g/ml}$  sheared denatured salmon sperm DNA at 60°C for 2 hours. Hybridization was carried out

overnight in 15 volume of the prehybridization solution with  $10^7$  CPM of probe added (final  $7 \times 10^5$  CPM/ml). Probes were hybridized to separate blots in parallel. Blots were washed in 5x SSC (0.75 M NaCl, 0.075 M NaCitrate) at 50°C for 30 minutes. Blots were exposed to a Packard MP Phosphor Screen for two hours and visualized using a Packard Cyclone™

5 Storage Phosphor System.

The results, shown in FIG. 4B, demonstrate that the probe labeled using the method of the present invention has a sensitivity in a dot blot hybridization assay about 10-fold greater than that of a 5'-end-labeled probe. Thus functional activity in a practical application directly parallels the predicted results based upon the measured specific activity of each

10 probe and directly demonstrates the utility of the invention.

### 7.5 NON-ISOTOPIC LABELING

Substrate Nucleic Acid (SEQ ID NO:17) and Template Nucleic Acid were labeled using the primer extension method of the invention as outlined above in Section 7.3.  $\alpha$ - $^{32}$ P-

15 dATP was replaced with biotin-7-dATP label at a final concentration of 10  $\mu$ M. The reaction proceeded normally and resulted in a highly modified biotin labeled probe. The biotin labeled probe was hybridized in a dot blot assay as described in Section 7.4 above, in which it successfully detected its target.

20 8. **EXAMPLE 3: COMPARISON OF SENSITIVITY OF PROBES  
LABELED BY PRIMER EXTENSION AND BY T4  
POLYNUCLEOTIDE KINASE IN A NORTHERN BLOT ASSAY**

#### 8.1 MATERIALS AND REAGENTS

Reagents and materials used to label the Substrate Nucleic Acid at the 5' end using T4 PNK and at the 3' end by primer extension were identical to the methods previously

25 described in Example 2 (Section 7, *supra*). In this example the Substrate Nucleic Acid, a 36-mer, and Template Nucleic Acid, a 16-mer, are again provided lyophilized. The Substrate Nucleic Acid contains a 30 base specific sequence at the 5' end complementary to the human  $\beta$ -actin mRNA and the standard 6-base 3' end Template Hybridization Domain. The Template Nucleic Acid is comprised of a 10-base oligo-dT Signal Template Domain

30 sequence at the 5'-end, a 6-base 3' Substrate Hybridization Domain complementary to the Template Hybridization Nucleic Acid Domain of the Substrate, and is blocked with a 3'-amino-modifier-C7 group.

The Substrate Nucleic Acid is:

5, - GCCCAGGAAGGAAGGCTGGAAGAGTGCCTCGGCGGG-3' (SEQ ID NO:20),

35 where the underlined portion is the Template Binding Domain. The Template Nucleic Acid is: 5'-TTTTTTTTTTCCCGCC-C7-amino-modifier-3' (SEQ ID NO:10), where C

represents C5-propynyl-dC.

The Template Nucleic Acid was resuspended in TE at a concentration of 12.5  $\mu\text{M}$  (12.5 picomoles per  $\mu\text{l}$ ). A stock solution of the Substrate Nucleic Acid was made in TE at a concentration of 10  $\mu\text{M}$  (10 picomoles per  $\mu\text{l}$ ). A working solution of the Substrate was made by 1:20 dilution in distilled water with a final concentration of 0.5  $\mu\text{M}$  (0.5 picomoles per  $\mu\text{l}$ ). All reagents were stored at  $-20^{\circ}\text{C}$  and kept in an ice bucket at  $0^{\circ}\text{C}$  until thawed for use. It is preferable to make a series of aliquots of these reagents upon receipt of the kit in single use tubes for both contamination control and ease of use.

## 8.2 LABELING WITH T4 POLYNUCLEOTIDE KINASE

From the 10  $\mu\text{M}$  stock of Substrate Nucleic Acid, 2 picomoles (2  $\mu\text{l}$ ) was 5'-end labeled using T4 polynucleotide kinase, as described in Example 2 (Section 7.2, *supra*). Substrate Nucleic Acid was incubated with 100  $\mu\text{Ci}$   $\gamma$ - $^{32}\text{P}$ -ATP and 5 units of T4 PNK in a final reaction volume of 50  $\mu\text{l}$  with kinase buffer for 1 hour at  $37^{\circ}\text{C}$ . Unincorporated  $\gamma$ - $^{32}\text{P}$ -ATP was removed by passage through a Sephadex G-25 spin column. Radiolabel incorporation was measured using scintigraphy and, assuming no loss of labeled nucleic acid or contamination with unincorporated label, resulted in a specific activity of  $8 \times 10^6$  CPM per picomole probe. Structure of this labeled probe is shown in FIG. 5A.

## 8.3 NUCLEIC ACID LABELING BY PRIMER EXTENSION

The annealing step of the primer extension reaction was carried out as described previously. Briefly, 1  $\mu\text{l}$  (0.5 picomoles) of Substrate Nucleic Acid (0.5  $\mu\text{M}$  stock), 1  $\mu\text{l}$  of 10X reaction buffer, and 1  $\mu\text{l}$  (12.5 picomoles) of Template Nucleic Acid were added together in a 500  $\mu\text{l}$  microcentrifuge tube. After heating the sample to  $94^{\circ}\text{C}$  (1 minute) and cooling to room temperature (5 minutes), 6  $\mu\text{l}$  of  $\alpha$ - $^{32}\text{P}$  dATP and 1  $\mu\text{l}$  (5 units) of the exo<sup>-</sup> Klenow fragment of *E. coli* DNA polymerase were added. The reaction was allowed to proceed at room temperature for one hour. 40  $\mu\text{l}$  of stop solution (10 mM EDTA) was added and the reaction was passed through a Sephadex G25 spin column to remove unincorporated  $\alpha$ - $^{32}\text{P}$ -dATP. An aliquot was removed and evaluated for activity by liquid scintillation counting. Assuming no loss of labeled nucleic acid or contamination with unincorporated label, the probe had a specific activity of  $7 \times 10^7$  CPM per picomole, about 10 times that of the probe labeled with PNK. The structure of the labeled probe produced is shown in FIG. 5A.

## 8.4 COMPARISON OF PROBE SENSITIVITIES IN A NORTHERN BLOT ASSAY

To directly compare the functional sensitivity of a PNK 5' end-labeled probe and a 3'

primer-extension labeled probe using the method of the invention, two identical Northern blots were prepared using human placental RNA. Ten micrograms of total RNA (Ambion) was separated on a formaldehyde-agarose gel (Current Protocols in Molecular Biology, John Wiley & Sons, Inc.). After electrophoresis, RNA was transferred to MSI nylon membranes (Micron Separations, Inc.) by capillary action in 20x SSC and baked at 80°C for 1 hour. The membranes were prehybridized in PerfectHyb™ (Sigma) at 60°C for one hour. Hybridization was carried out for 4 hours at 60°C in 10 mls total volume of PerfectHyb™ with 10<sup>7</sup> CPM of probe added (final 10<sup>6</sup> CPM/ml). Probes were hybridized to separate blots in parallel. Blots were washed twice in 5x SSC (0.75 M NaCl, 0.075 M NaCitrate) at 55°C for 30 minutes. Blots were exposed to a Packard MP Phosphor Screen for one hour and visualized using a Packard Cyclone™ Storage Phosphor System.

The results, shown in FIG. 5B, demonstrate that the probe labeled using the method of the present invention has greater sensitivity in a Northern blot hybridization assay than a 5'-end-labeled probe. Thus functional activity in a second application shows improved sensitivity using a high specific activity probe and directly demonstrates the utility of the invention.

## 9. EXAMPLE 4: USE OF THE PRIMER EXTENSION LABELING METHOD WITH AN RNA SUBSTRATE NUCLEIC ACID

### 9.1 MATERIALS AND REAGENTS

Reagents and methods used to label an oligoribonucleotide at the 5'-end with PNK were identical to those used to label oligodeoxyribonucleotides in Example 1 (Section 6, *supra*).

In this example, two Substrate Nucleic Acids, one DNA and one RNA, were labeled by the method of the invention. The oligodeoxyribonucleotide employed was 5'-GCCCAGGAAGGAAGGCTGGAAGAGTGCCTCGGCGGG-3' (SEQ ID NO:20), a 36-mer comprised of a unique 30 nucleotide sequence at the 5'-end and a 6-base 3' Template Hybridization Domain (underlined). The oligoribonucleotide employed was 5'-cugggcauggaguccuguggcauccacgaacuaccuacaggggg-3' (SEQ ID NO:23), a 46-mer comprised of a unique 40 nucleotide sequence at the 5'-end and a 6-base 3' Template Hybridization Domain (underlined). One oligodeoxyribonucleotide Template Nucleic Acid was employed for both reactions, 5'-TTTTTTTTTTPcCpCgCpC-C<sub>7</sub>-amino-modifier-3' (SEQ ID NO:10).

To obtain a quantitative visualization of Substrate utilization, radiolabel was incorporated at the 5'-end of the Substrate Nucleic Acids using T4 PNK and γ-<sup>32</sup>P-ATP. The subsequent primer extension reaction was conducted using unlabeled (cold) dATP, as described in Example 1 (Section 6 above).

## 9.2 LABELING NUCLEIC ACIDS USING T4 KINASE

The Substrate Nucleic Acids were resuspended in TE at a final concentration of 0.5  $\mu$ M (0.5 picomoles/ $\mu$ l). From this stock, 2 picomoles (4  $\mu$ l) was 5' end labeled using T4  
5 PNK. Substrate Nucleic Acids were each incubated with 100  $\mu$ Ci  $\gamma$ - $^{32}$ P-ATP and 5 units of T4 PNK in a final reaction volume of 50  $\mu$ l with kinase buffer for 1 hour at 37°C. Unincorporated  $\gamma$ - $^{32}$ P-ATP was removed by passage through a Sephadex G-25 spin column. Radiolabel incorporation was measured using scintigraphy.

The Template Nucleic Acid was resuspended at 12.5  $\mu$ M concentration in TE (12.5  
10 picomoles per  $\mu$ l). 5'- $^{32}$ P-end-labeled Substrate Nucleic Acids (prepared as described above) were "spiked" into aliquots of cold Substrate at 10,000 CPM/ $\mu$ l (approximately 1 femtomole of radiolabeled compound per 0.5 picomole cold oligonucleotide). All reagents were stored at -20°C and kept in an ice bucket at 0°C until thawed for use.

## 15 9.3 PRIMER EXTENSION REACTION

The primer extension protocol involved two steps, annealing the Substrate and Template Nucleic Acids and extending the Substrate Nucleic Acid in a primer extension reaction. Following the extension reaction, the reaction products were separated using analytical PAGE followed by visualization using phosphorimaging.

20 1  $\mu$ l of Substrate Nucleic Acid, 1  $\mu$ l of 10x stock reaction buffer, and 1  $\mu$ l of Template Nucleic Acid were added together in a 500  $\mu$ l microcentrifuge tube in the molar concentrations shown in Table 2 (see FIG. 1B). The reaction was mixed without vortexing by pipetting up and down. The tube was then placed in a hot water bath at 94°C for 1 minute. The tube was removed from heat and the annealing reaction proceeded in the same  
25 tube for 5 minutes, slowly cooling to room temperature. The tube was then centrifuged briefly to return the contents to the bottom of the tube, which ensured that all nucleic acid and buffer was available for the primer extension reaction.

1  $\mu$ l of 10 $\mu$ M unlabeled dATP, 5  $\mu$ l of water, and 1  $\mu$ l of exo- Klenow fragment of *E. coli* DNA polymerase I were added directly into the previously prepared annealing tube  
30 and mixed using an up and down pipetting action with an aerosol barrier pipette tip. The reaction was allowed to proceed at room temperature for one hour (FIG. 1C). The reaction was stopped with the addition of an equal volume (10  $\mu$ l) of gel loading buffer (7M urea, 100 mM EDTA), heated at 94°C for 5 minutes, and separated on a 12% polyacrylamide, 7M urea, 0.5x TBE (45 mM Tris borate, 1 mM EDTA) gel. The gel was exposed to a Packard  
35 MP Phosphor Screen for 2 hours and visualized using a Packard Cyclone™ Storage Phosphor System. In this example, the 5'-end labeled 36-mer DNA Substrate Nucleic Acid



5'-<sup>32</sup>P-GCCCAGGAAGGAAGGCTGGAAGAGTGCCTCGGCGGG-3' (SEQ ID NO:21)  
is increased in length to 46 bases,

5'-<sup>32</sup>P-GCCCAGGAAGGAAGGCTGGAAGAGTGCCTCGGCGGGAAAAAAAAA-3'  
(SEQ ID NO:22). The 5'-end labeled 46-mer RNA Substrate Nucleic Acid,

- 5 5'-<sup>32</sup>P-cugggcaggaguccuguggcauccacgaaacuaccuucaggcggg-3' (SEQ ID NO:24), is  
increased in length to 56 bases by the primer extension reaction to

5'-<sup>32</sup>P-cugggcaggaguccuguggcauccacgaaacuaccuucaggcgggAAAAAAAAA-3' (SEQ ID  
NO:25), as shown in FIG. 6A. The extension efficiency was determined by comparing the  
amount of signal visualized in each product by phosphorimaging, as shown in FIG. 6B.

- 10 These values were quantitatively determined by integrating pixel density (ImageQuant™  
Software, Packard Instrument Company) from images obtained using a Packard Cyclone™  
Storage Phosphor System. Approximately 90% of the DNA substrate was extended in the  
primer extension reaction while just over 70% of the RNA substrate was extended in the  
primer extension reaction. The methods described herein, therefore, can be used to label  
15 RNA Substrate Nucleic Acids.

- The present invention is not to be limited in scope by the specific embodiments  
described herein. Indeed, various modifications of the invention in addition to those  
described herein will become apparent to those skilled in the art from the foregoing  
20 description and accompanying figures. Such modifications are intended to fall within the  
scope of the appended claims.

Various publications are cited herein, the disclosures of which are incorporated by  
reference in their entireties.

25

30

35

What is claimed is:

1. A Complex, comprising:

- a. a first nucleic acid comprising, from 3' to 5': a Substrate Hybridization Domain and a Signal Template Domain, wherein:
- i. the Substrate Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides; and
  - ii. the Signal Template Domain comprises a sequence of about 5 to about 100 nucleotides;

and:

- b. a second nucleic acid comprising from 3' to 5': a Signal Domain, a Template Hybridization Domain and a Target Binding Domain, wherein:
- i. the Signal Domain comprises a sequence of about 5 to about 100 nucleotides, which sequence shows complementarity toward and is hybridizable to the Signal Template Domain of the first nucleic acid, and of which at least two nucleotides are detectably labeled;
  - ii. the Template Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides, is not detectably labeled, and shows complementarity toward and is hybridizable to the Substrate Hybridization Domain of the first nucleic acid;
  - iii. the Target Binding Domain is not detectably labeled and comprises a nucleotide sequence heterologous to that of the Template Hybridization Domain.

2. The Complex of claim 1, wherein the nucleotides which comprise the first or second nucleic acid are deoxyribonucleotides.

3. The Complex of claim 1, wherein the nucleotides which comprise the Signal Domain of the second nucleic acid are deoxyribonucleotides and the nucleotides which comprise the Template Hybridization Domain and the Target Binding Domain of the second nucleic acid are ribonucleotides.

4. The Complex of claim 1, wherein the Substrate Hybridization Domain is at the 3' end of the first nucleic acid.

5. The Complex of claim 1, wherein the Substrate Hybridization Domain comprises a sequence of about 5 to about 10 nucleotides.

5 6. The Complex of claim 1, wherein the Substrate Hybridization Domain cannot be extended by a 5'-3' DNA polymerase.

7. The Complex of claim 6, wherein the Substrate Hybridization Domain further comprises an extension of nucleotides at the 3' end of said Substrate Hybridization  
10 Domain, the extension having no complementarity to the Template Hybridization Domain of the second nucleic acid.

8. The Complex of claim 6, wherein the Substrate Hybridization Domain comprises a 3'-terminal modified nucleotide.

15 9. The Complex of claim 8, wherein the modification is selected from the group consisting of: a 3'-amino-modifier, a 2', 3'-dideoxynucleotide, a 3'-phosphate, and a modified 3'-phosphate group.

20 10. The Complex of claim 1, wherein the Substrate Hybridization Domain comprises at least one nucleotide which comprises a modified cytidine, which nucleotide is selected from the group consisting of: C5-methyl-dC and C5-propynyl-dC.

11. The Complex of claim 1, wherein the Signal Domain comprises a sequence  
25 of about 10 to about 50 nucleotides.

12. The Complex of claim 1, wherein the Signal Domain is at least 50%, at least 70%, at least 90% or 100% homopolymeric.

30 13. The Complex of claim 1, wherein at least 60%, at least 80% or 100% of the nucleotides of the Template Hybridization Domain comprise guanosine or cytidine or a combination thereof.

14. The Complex of claim 1, wherein at least 60% of the nucleotides of the  
35 Template Hybridization Domain comprise guanosine or cytidine or a combination thereof, and the Signal Domain is at least 50% homopolymeric.

15. A reaction mixture for use in a process for the labeling of a nucleic acid molecule comprising:

- 5           a. a first nucleic acid comprising, from 3' to 5': a Substrate Hybridization Domain and a Signal Template Domain, wherein:
- i. the Substrate Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides; and
- ii. the Signal Template Domain comprises a sequence of about 5 to about 100 nucleotides;
- 10           and:
- b. a second nucleic acid comprising from 3' to 5': a Template Hybridization Domain and a Target Binding Domain, wherein:
- i. the Template Hybridization Domain comprises a sequence of about 5 to about 20 nucleotides, is not detectably labeled, and
- 15           shows complementarity toward and is hybridizable to the Substrate Hybridization Domain of the first nucleic acid;
- ii. the Target Binding Domain is not detectably labeled and comprises a nucleotide sequence heterologous to that of the Template Hybridization Domain.

20

16. The reaction mixture of claim 15, wherein the nucleotides which comprise the first or second nucleic acid are deoxyribonucleotides.

17. The reaction mixture of claim 15, wherein the nucleotides which comprise 25 the first or second nucleic acid are ribonucleotides.

18. The reaction mixture of claim 15, wherein the Substrate Hybridization Domain is at the 3' end of the first nucleic acid.

30       19. The reaction mixture of claim 15, wherein the Substrate Hybridization Domain comprises a sequence of about 5 to about 10 nucleotides.

20. The reaction mixture of claim 15, wherein the Substrate Hybridization Domain cannot be extended by a 5'→3' DNA polymerase.

35

21. The reaction mixture of claim 20, wherein the Substrate Hybridization

Domain further comprises an extension of nucleotides at the 3' end of said Substrate Hybridization Domain, the extension having no complementarity to the Template Hybridization Domain of the second nucleic acid.

5           22.    The reaction mixture of claim 20, wherein the Substrate Hybridization Domain comprises a 3'-terminal modified nucleotide.

          23.    The reaction mixture of claim 22, wherein the modification is selected from the group consisting of: a 3'-amino-modifier, a 2', 3'-dideoxynucleotide, a 3'-phosphate, and  
10 a modified 3'-phosphate group.

          24.    The reaction mixture of claim 15, wherein the Substrate Hybridization Domain comprises at least one nucleotide which comprises a modified cytidine, which nucleotide is selected from the group consisting of: C5-methyl-dC and C5-propynyl-dC.

15

          25.    The reaction mixture of claim 15, wherein the Signal Template Domain comprises a sequence of about 10 to about 50 nucleotides.

          26.    The reaction mixture of claim 15, wherein the Signal Domain is at least 50%,  
20 at least 70%, at least 90% or 100% homopolymeric.

          27.    The reaction mixture of claim 15, wherein at least 60%, at least 80% or 100% of the nucleotides of the Template Hybridization Domain comprise guanosine or cytidine or a combination thereof.

25

          28.    The reaction mixture of claim 15, wherein at least 60% of the nucleotides of the Template Hybridization Domain comprise guanosine or cytidine or a combination thereof, and the Signal Domain is at least 50% homopolymeric.

30           29.    A method of labeling a nucleic acid molecule, comprising the steps of:

- a.    Hybridizing a first nucleic acid to a second nucleic acid, wherein the first nucleic acid comprises, from 3' to 5': a Substrate Hybridization Domain and a Signal Template Domain, wherein:
  - i.    the Substrate Hybridization Domain comprises a sequence of  
35 about 5 to about 20 nucleotides; and
  - ii.   the Signal Template Domain comprises a sequence of about 5

to about 100 nucleotides;  
and the second nucleic acid comprises from 3' to 5': a Template Hybridization Domain and a Target Binding Domain, wherein:  
i. the Template Hybridization Domain comprises a sequence of  
5 about 5 to about 20 nucleotides, is not detectably labeled, and shows complementarity toward and is hybridizable to the Substrate Hybridization Domain of the first nucleic acid;  
ii. the Target Binding Domain is not detectably labeled and comprises a nucleotide sequence heterologous to that of the  
10 Template Hybridization Domain;

and:

b. extending the second nucleic acid with a DNA polymerase in the presence of a labeled nucleotide to create a Signal Domain having a sequence which shows complementarity toward and is hybridizable  
15 to the Signal Template Domain, thereby labeling said second nucleic acid molecule.

30. The method of claim 29, wherein the nucleotides which comprise the first or second nucleic acid are deoxyribonucleotides.

31. The method of claim 29, wherein the nucleotides which comprise the first or second nucleic acid are ribonucleotides.

32. The method of claim 29, wherein the second nucleic acid consists of about  
25 15 to about 150 nucleotides.

33. The method of claim 29, wherein the Substrate Hybridization Domain is at the 3' end of the first nucleic acid.

34. The method of claim 29, wherein the Substrate Hybridization Domain comprises a sequence of about 5 to about 10 nucleotides.

35. The method of claim 29, wherein the Substrate Hybridization Domain cannot be extended by a 5'→3' DNA polymerase.

36. The method of claim 35, wherein the Substrate Hybridization Domain

further comprises an extension of nucleotides at the 3' end of said Substrate Hybridization Domain, the extension having no complementarity to the Template Hybridization Domain of the second nucleic acid.

5           37.     The method of claim 35, wherein the Substrate Hybridization Domain comprises a 3'-terminal modified nucleotide.

            38.     The method of claim 37, wherein the modification is selected from the group consisting of: a 3'-amino-modifier, a 2', 3'-dideoxynucleotide, a 3'-phosphate, and a  
10 modified 3'-phosphate group.

            39.     The method of claim 29, wherein the Substrate Hybridization Domain comprises at least one nucleotide which comprises a modified cytidine, which nucleotide is selected from the group consisting of: C5-methyl-dC and C5-propynyl-dC.

15

            40.     The method of claim 29, wherein the Signal Template Domain comprises a sequence of about 10 to about 50 nucleotides.

            41.     The method of claim 29, wherein the Signal Domain is at least 50%, at least  
20 70%, at least 90% or 100% homopolymeric.

            42.     The method of claim 29, wherein at least 60%, at least 80% or 100% of the nucleotides of the Template Hybridization Domain comprise guanosine or cytidine or a combination thereof.

25

            43.     The method of claim 29, wherein at least 60% of the nucleotides of the Template Hybridization Domain comprise guanosine or cytidine or a combination thereof, and the Signal Domain is at least 50% homopolymeric.

30           44.     The method of claim 29, wherein the extending step is carried out by a DNA polymerase selected from the group consisting of: *E. coli* DNA polymerase I holoenzyme, Klenow fragment of *E. coli* DNA polymerase I, T4 DNA polymerase, T7 DNA polymerase, and a DNA polymerase encoded by a thermophilic bacterium.

35           45.     The method of claim 29, wherein the Template Hybridization Domain or the Substrate Hybridization Domain comprises at least one modified nucleotide, which

modified nucleotide increases the hybridization affinity of said Template Hybridization Domain to said Substrate Hybridization Domain.

46. The method of claim 45, wherein at least one modified nucleotide is found in  
5 the Template Hybridization Domain.

47. The method of claim 46, wherein at least one modified nucleotide is selected from the group consisting of: C5-methyl-dC, C5-propynyl-dC, C5-propynyl-dU, and 2, 6-diaminopurine.

10

48. The method of claim 29, wherein at least one nucleotide comprises a label selected from the group consisting of:  $^{32}\text{P}$ ,  $^{33}\text{P}$ ,  $^{35}\text{S}$ , fluorescein, digoxigenin, biotin, Cy5, Cy3, and rhodamine.

15

49. A method for detecting a Target Nucleic Acid in a sample, comprising:

- a. contacting the sample with the Complex of claim 1 under conditions whereby said Complex can bind to the Target Nucleic Acid to form a Complex-Target Nucleic Acid hybrid; and
- b. detecting any Complex-Target Nucleic Acid hybrids, so that if a  
20 Complex-Target Nucleic Acid hybrid is detected, a Target Nucleic Acid is detected in the sample.

20

50. A method for detecting a Target Nucleic Acid in a sample, comprising:

- a. dissociating the Complex of claim 1 to generate a single stranded first nucleic acid and a single stranded second nucleic acid;
- b. contacting the sample with the second nucleic acid of step a. under conditions whereby said second nucleic acid can bind to the Target Nucleic Acid to form a second nucleic acid-Target Nucleic Acid hybrid; and
- c. detecting any second nucleic acid-Target Nucleic Acid hybrids, so  
30 that if a second nucleic acid-Target Nucleic Acid hybrid is detected, a Target Nucleic Acid is detected in the sample.

25

30

51. A kit for labeling a nucleic acid molecule, comprising the reaction mixture of  
35 claim 15, and a DNA polymerase.



52. The kit of claim 51, wherein at least 60% of the nucleotides of the Template Hybridization Domain comprise guanosine or cytidine or a combination thereof, and the Signal Domain is at least 50% homopolymeric.

5 53. The kit of claim 51, wherein the Substrate Hybridization Domain comprises a predetermined sequence comprising CCCGCC and the Signal Template Domain comprises a predetermined sequence comprising TTTTTTTTTT.

54. The kit of claim 51, wherein, the first nucleic acid comprises a  
10 predetermined sequence comprising SEQ ID NO:10.

15

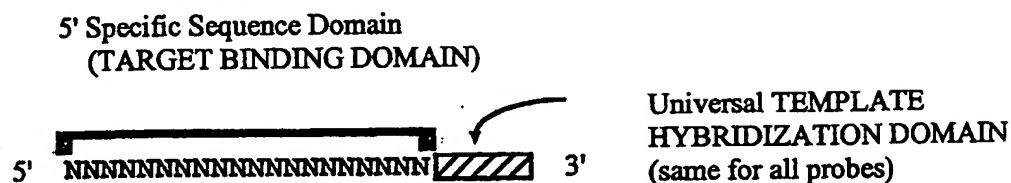
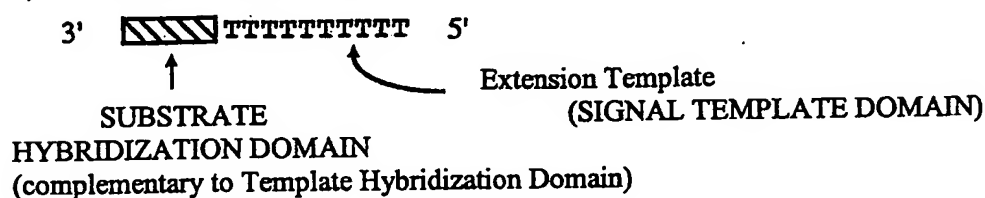
20

25

30

35

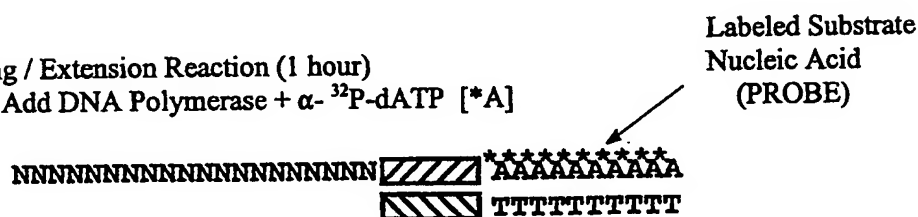
## A) Design and Synthesis of Nucleic Acids (oligonucleotides):

i) *Substrate Nucleic Acid*ii) *Template Nucleic Acid*

## B) Anneal Substrate + Template Nucleic Acids in Buffer (10 minutes)



## C) Labeling / Extension Reaction (1 hour)

Add DNA Polymerase +  $\alpha$ -<sup>32</sup>P-dATP [\*A]

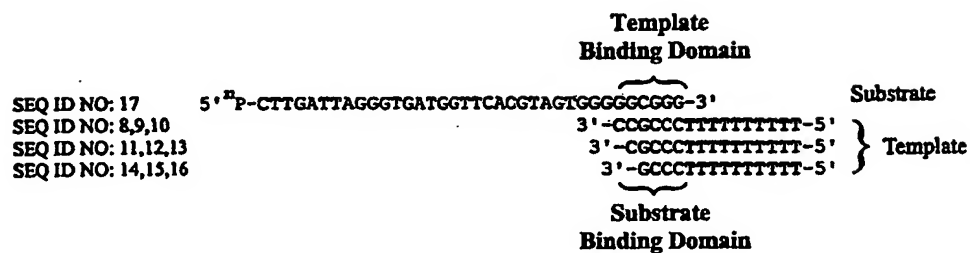
## D) Remove Unincorporated Label using Column Chromatography (10 minutes)

Optional: Remove Template and more highly purify Probe using PAGE (2 hour gel + elute)



FIG. 1

## A) Substrate and Template Nucleic Acids



## B) PAGE analysis of reaction products

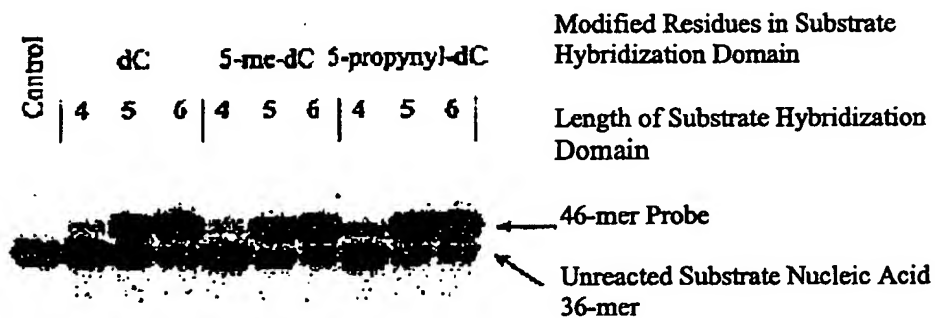


FIG. 2

BEST AVAILABLE COPY

BEST AVAILABLE COPY

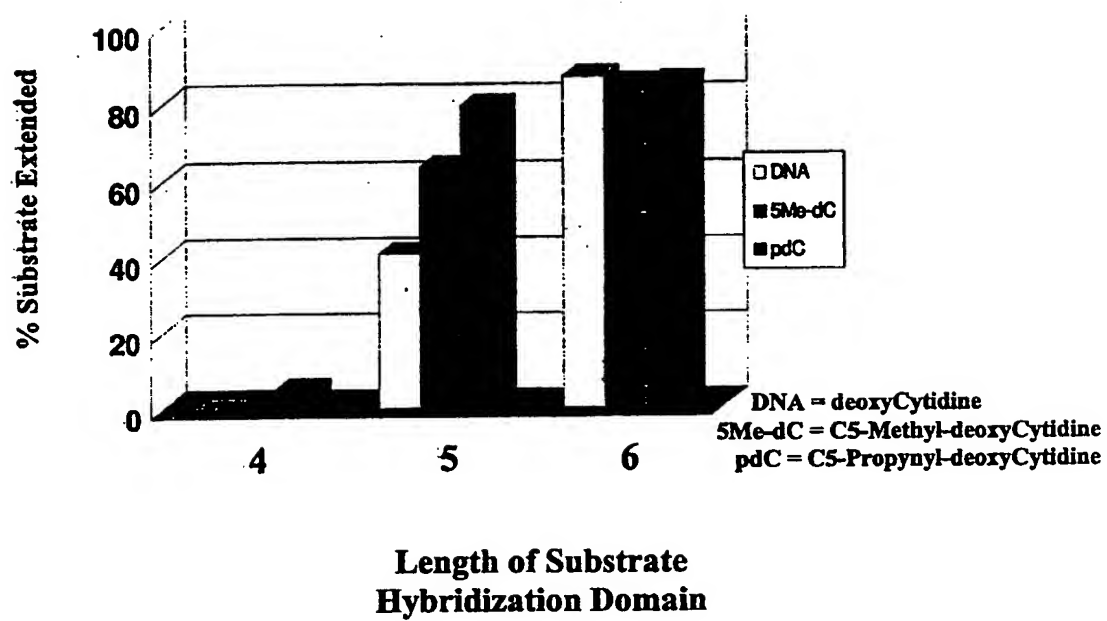


FIG. 3

**pGreen Lantern™-1 Probes****Kinase Labeled Probe**

5' <sup>32</sup>P-CTTGATTAGGGTGATGGTTCACGTAGTGGGGGCGGG-3'

**High Specific Activity Tailed Probe**

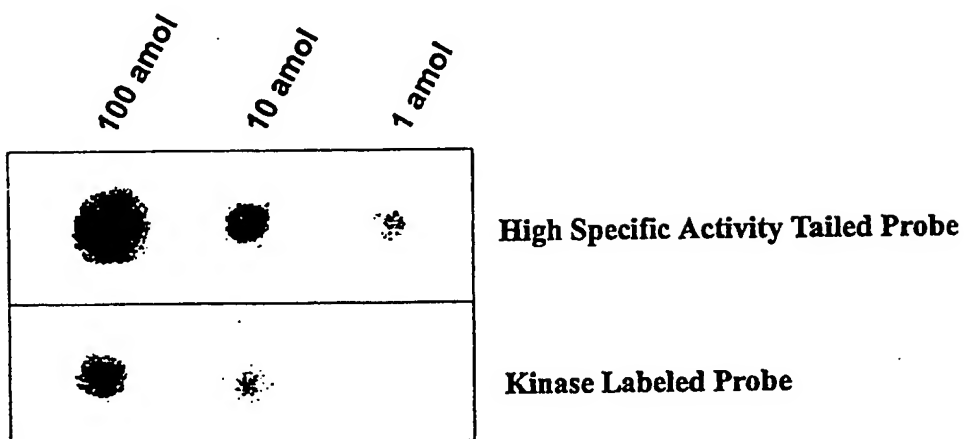
5'-CTTGATTAGGGTGATGGTTCACGTAGTGGGGGCGGG\*A\*A\*A\*A\*A\*A\*A\*A\*A\*A-3'

↑  
\* = <sup>32</sup>P (10 total)

**FIG. 4A**

**Dot Blot**

Mass of target plasmid on Membrane



BEST AVAILABLE COPY

**FIG. 4B**

**$\beta$ -Actin Probes****Kinase Labeled Probe**

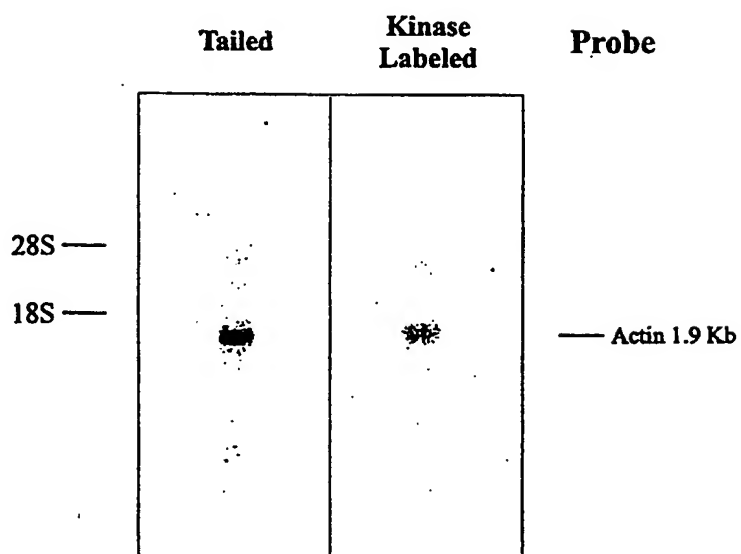
5' <sup>32</sup>P-GCCCAGGAAGGAAGGCTGGAAGAGTGCCTCGGCGGG-3'

**High Specific Activity Tailed Probe**

5'-GCCCAGGAAGGAAGGCTGGAAGAGTGCCTCGGCGGG\*A\*A\*A\*A\*A\*A\*A\*A\*A\*A-3'

↑  
\* = <sup>32</sup>P (10 total)

**FIG. 5A**

**Northern Blot of Human Placental RNA**

BEST AVAILABLE COPY

**FIG. 5B**



## Substrate Nucleic Acids and Reaction Products

SEQ ID NO: 21 DNA Substrate Nucleic Acid

5' <sup>32</sup>P-GCCCAGGAAGGAAGGCTGGAAGAGTGCCTCGGCGGG-3'

SEQ ID NO: 22 DNA Reaction Product

5' <sup>32</sup>P-GCCCAGGAAGGAAGGCTGGAAGAGTGCCTCGGCGGGAAAAAAAAA-3'

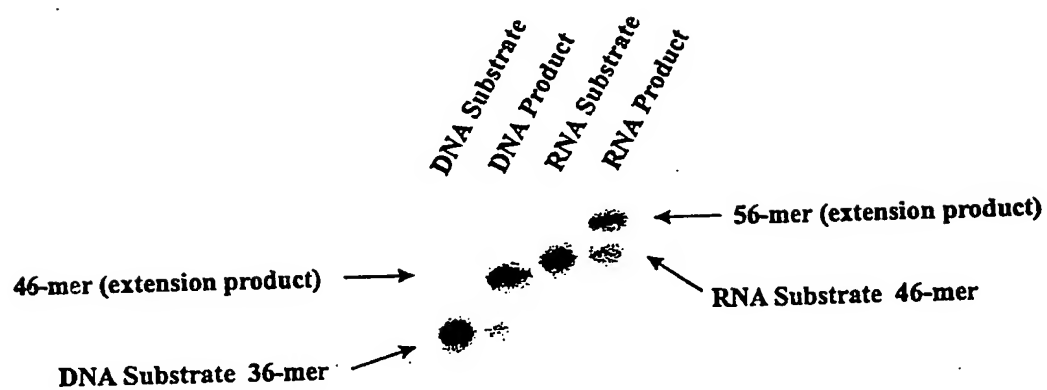
SEQ ID NO: 24 RNA Substrate Nucleic Acid

5' <sup>32</sup>P-cugggcauggaguccuguggcauccacgaaacuaccuucaggcggg-3'

SEQ ID NO: 25 RNA Reaction Product

5' <sup>32</sup>P-cugggcauggaguccuguggcauccacgaaacuaccuucaggcgggAAAAAAAAA-3'

**FIG. 6A**

**PAGE analysis of Reaction Products**

BEST AVAILABLE COPY

**FIG. 6B**

## SEQUENCE LISTING

<110> INTEGRATED DNA TECHNOLOGIES, INC.

<120> PRIMER EXTENSION METHODS FOR PRODUCTION OF HIGH  
SPECIFIC ACTIVITY NUCLEIC ACID PROBES

<130> 7614-018

<140> To Be Assigned

<141> 2000-02-04

<160> 25

<170> PatentIn Ver. 2.1

<210> 1

<211> 6

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Artificial  
Sequence

<400> 1

ggcggg

6

<210> 2

<211> 6

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Artificial  
Sequence

<400> 2

gggggg

6

<210> 3

<211> 6

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Artificial  
Sequence

<400> 3

cccccc

6

<210> 4

<211> 20

<212> DNA  
<213> Artificial Sequence

<220>  
<221> modified\_base  
<222> (1)..(20)  
<223> N = A, C, G, or T

<220>  
<223> Description of Artificial Sequence: Artificial  
Sequence

<400> 4  
nnnnnnnnnn nnnnnnnnnn

20

<210> 5  
<211> 26  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> modified\_base  
<222> (1)..(20)  
<223> N = A, C, G, or T

<220>  
<223> Description of Artificial Sequence: Artificial  
Sequence

<400> 5  
nnnnnnnnnn nnnnnnnnnn ggcggg

26

<210> 6  
<211> 10  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Artificial  
Sequence

<400> 6  
tttttttttt

10

<210> 7  
<211> 7  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> modified\_base  
<222> (7)  
<223> N = C7-AMINO-MODIFIER, 3'

<220>

<223> Description of Artificial Sequence: Artificial Sequence

<400> 7  
cccgccn

7

<210> 8  
<211> 17  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> modified\_base  
<222> (17)  
<223> N = C7-AMINO-MODIFIER, 3'

<220>  
<223> Description of Artificial Sequence: Artificial Sequence

<400> 8  
ttttttttt ccgccc

17

<210> 9  
<211> 17  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> modified\_base  
<222> (11)..(13)  
<223> N = 5-METHYL-2'-DEOXYCYTOSINE

<220>  
<221> modified\_base  
<222> (15)..(16)  
<223> N = 5-METHYL-2'-DEOXYCYTOSINE

<220>  
<221> modified\_base  
<222> (17)  
<223> N = C7-AMINO-MODIFIER, 3'

<220>  
<223> Description of Artificial Sequence: Artificial Sequence

<400> 9  
ttttttttt nnnnnnn

17

<210> 10  
<211> 17  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> modified\_base  
<222> (11)..(13)  
<223> N= 5-PROPYNYL-2'-DEOXYCYTOSINE  
  
<220>  
<221> modified\_base  
<222> (15)..(16)  
<223> N = 5-PROPYNYL-2'-DEOXYCYTOSINE  
  
<220>  
<221> modified\_base  
<222> (17)  
<223> N = C7-AMINO-MODIFIER  
  
<220>  
<223> Description of Artificial Sequence: Artificial  
Sequence

<400> 10  
tttttttttt nnnngnnn

17

<210> 11  
<211> 16  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> modified\_base  
<222> (16)  
<223> N = C7-AMINO-MODIFIER, 3'

<220>  
<223> Description of Artificial Sequence: Artificial  
Sequence

<400> 11  
tttttttttt cccgcn

16

<210> 12  
<211> 16  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> modified\_base  
<222> (11)..(13)  
<223> N = 5-METHYL-2'-DEOXYCYTOSINE

<220>  
<221> modified\_base  
<222> (15)  
<223> N = 5-METHYL-2'-DEOXYCYTOSINE

<220>  
<221> modified\_base

<222> (16)  
<223> N = C7-AMINO-MODIFIER, 3'

<220>  
<223> Description of Artificial Sequence: Artificial  
Sequence

<400> 12  
tttttttttt nnnngnn

16

<210> 13  
<211> 16  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> modified\_base  
<222> (11)..(13)  
<223> N = 5-PROPYNYL-2'-DEOXYCYTOSINE

<220>  
<221> modified\_base  
<222> (15)  
<223> N = 5-METHYL-2'-DEOXYCYTOSINE

<220>  
<221> modified\_base  
<222> (16)  
<223> N = C7-AMINO-MODIFIER

<220>  
<223> Description of Artificial Sequence: Artificial  
Sequence

<400> 13  
tttttttttt nnnngnn

16

<210> 14  
<211> 15  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> modified\_base  
<222> (15)  
<223> N = C7 AMINO-MODIFIER, 3'

<220>  
<223> Description of Artificial Sequence: Artificial  
Sequence

<400> 14  
tttttttttt cccggn

15

<210> 15

<211> 15  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> modified\_base  
<222> (15)  
<223> N = C7-AMINO-MODIFIER, 3'

<220>  
<221> modified\_base  
<222> (11)..(13)  
<223> N = 5-METHYL-2'-DEOXYCYTOSINE

<220>  
<223> Description of Artificial Sequence: Artificial  
Sequence

<400> 15  
tttttttttt nnngn

15

<210> 16  
<211> 15  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> modified\_base  
<222> (11)..(13)  
<223> N = 5-PROPYNYL-2'-DEOXYCYTOSINE

<220>  
<221> modified\_base  
<222> (15)  
<223> N = C7-AMINO-MODIFIER, 3'

<220>  
<223> Description of Artificial Sequence: Artificial  
Sequence

<400> 16  
tttttttttt nnngn

15

<210> 17  
<211> 36  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Artificial  
Sequence

<400> 17  
cttgattagg gtgatgggtc acgtagtggg ggcggg

36



<210> 18  
<211> 36  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Artificial  
Sequence

<400> 18  
cttgattagg gtgatgggttc acgtagtggg ggcggg

36

<210> 19  
<211> 46  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Artificial  
Sequence

<400> 19  
cttgattagg gtgatgggttc acgtagtggg ggcgggaaaa aaaaaa

46

<210> 20  
<211> 36  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Artificial  
Sequence

<400> 20  
gcccaggaag gaaggctgga agagtgcctc ggcggg

36

<210> 21  
<211> 36  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Artificial  
Sequence

<400> 21  
gcccaggaag gaaggctgga agagtgcctc ggcggg

36

<210> 22  
<211> 46  
<212> DNA  
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Artificial  
Sequence

<400> 22

gcccaggaag gaaggctgga agagtgcctc ggcgggaaaa aaaaaa

46

<210> 23

<211> 46

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Artificial  
Sequence

<400> 23

cugggcaugg aguccugugg cauccacgaa acuaccuua ggcggg

46

<210> 24

<211> 46

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Artificial  
Sequence

<400> 24

cugggcaugg aguccugugg cauccacgaa acuaccuua ggcggg

46

<210> 25

<211> 55

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Artificial  
Sequence

<400> 25

cugggcaugg aguccugugg cauccacgaa acuaccuua ggcgggaaaa aaaaa

55

## INTERNATIONAL SEARCH REPORT

 International application No.  
 PCT/US00/02897

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> IPC(7) : CO7H 21/02, 19/04 US CL : 536/ 23.1, 24.3, 24.33, 26.6 According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) U.S. : 536/ 23.1, 24.3, 24.33, 26.6 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) Please See Extra Sheet.		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X — Y	US 5,538,872 A (BAHL et al) 23 July 1996(23.07.96), see entire document.	1-6, 11, 15-20, 25 — 29-50
Y	US 4,808,520 A (DATTA GUPTA et al) 28 February 1989(28.02.89), see entire document.	15-50
A	SHCHEPINOV et al. Oligonucleotide dendrimers: Synthesis and use as polylabelled DNA probes. Nucleic Acids Research. December 1997. Vol. 25, No. 22, pages 4447-4454, see entire document.	1-7, 15-21, 29-36, 44, 48-53
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "A" document member of the same patent family		
Date of the actual completion of the international search 19 MAY 2000		Date of mailing of the international search report 21 JUN 2000
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230		Authorized officer <i>Cynthia Wilder</i> Cynthia Wilder Telephone No. (703) 308-0196

**INTERNATIONAL SEARCH REPORT**International application No.  
PCT/US00/02897**C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MAAG et al. Oligonucleotide probes with multiple labels linked to the 4'-position of thymidine monomers: Excellent duplex stability and detection sensitivity. Tetrahedron Letters. August 1994. Vol. 35, No. 35, pages 6449-6452, see entire document.	1-53

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/02897

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2. ☒ Claims Nos.: 54  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:  
  
The CRF for this case is defective. Therefore, claim 54 drawn to a sequence comprising SEQ ID NO:10 is unsearchable.
  
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.  
☐ No protest accompanied the payment of additional search fees.

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/02897

### B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

Medline, embase, biotechds, biosis, scisearch, cancerlit, West 2.0, Dialog -Dissertation Abstract  
search terms: probe(s), primer extension, primer, label, labeling, multiple labels, hybridization, domain, substrate,  
template, polymerase, DNA, modified and cytidine, homopolymeric, fluorescein, biotin, radiolabel, polynucleotide kinase